

---

# Towards Qualitative Models of Interactions in Evolutionary Algorithms

---

**Karsten Weicker and Nicole Weicker**

Institute for Formal Methods of Computer Science

University of Stuttgart, P.O. Box 80 11 40

70511 Stuttgart, Germany

{Karsten,Nicole}.Weicker@informatik.uni-stuttgart.de

## Abstract

Most of today's theory and empirical work cannot serve as a foundation for designing or applying evolutionary algorithms in novel areas of application. This article proposes qualitative models of interactions in evolutionary algorithms as a means to close that gap. It presents a knowledge representation in the form of a hypergraph and exemplifies the integration of empirical results into the model by an examination of contracting recombination operators with respect to the diversity. Both the use of a big unifying model and small focused models are discussed. In the long term, this model is aimed to support the design, application, and debugging of evolutionary algorithms.

## 1 INTRODUCTION AND MOTIVATION

There exists a gap between what theoreticians provide and what practitioners need as a foundation for developing new problem-specific evolutionary algorithms (EA). Why is this? The theoretical results often do not suit the special problem at hand. And the existing experimental findings, that may be helpful, have to be extracted from thousands of publications since no comprehensive representation of the experimental findings exists.

There exist only few design rules or design philosophies (e.g. Radcliffe and Surry, 1995). Also simple recipes, rules of thumb, and checklists are scarce or even not useful at all. Therefore, this paper proposes the usage of qualitative models to collect the manifold

interactions and dependencies within the processing of evolutionary computing. Similar models are broadly used in development and diagnosis in the field of engineering (Rajagopalan, 1986).

The qualitative model aims at supporting the development and the fault analysis of evolutionary algorithms. Knowledge is represented as qualitative consequence of changes in certain parameters or characteristics. Both theoretical results and empirical knowledge may enter the model. In the long run, we anticipate to answer questions like by qualitative model-based reasoning:

*Given an algorithm  $X$  (with certain parameter values) and a goal  $Y$ . What are the consequences for  $Y$  of changing a special parameter (e.g. doubling the mutation rate)?*

It is a tool to get insight into the typical or presumable behavior in evolutionary computing. We also plan to use it as an educational resource.

In this paper we propose the *network of effects and dependencies* as a suitable representation for gathering knowledge on the internal dynamics of EAs. It is defined as a hypergraph where the chosen operators and parameters as well as the goals are external nodes in the sense that they do not depend on the evolutionary processing. The interaction of the components of the EA is modeled by internal nodes (keystones) and hyperedges. The keystones represent certain characteristics of the evolutionary processing that cannot be influenced directly by one parameter of the EA. The directed and labeled hyperedges describe the dependencies between the different external and internal nodes. The hypergraph can be transformed into a description using first order logic where each hyperedge equals a causal rule.

The pattern of describing the evolutionary dynamics by a model using keystones and rather simple interactions between the keystones is taken from artificial life where complex behavior stems from simple agents and simple interactions too. EAs are described by the observation of simple subprocesses and properties.

In many theoretical results the evolutionary dynamics are described by a single closed formula—usually on a macroscopic level (Rudolph, 1998; Poli, 1999; Shapiro, 2001). Usually the analysis using a macroscopic level is restricted to very simple problems or algorithms. Such a global approach with a general formula in mind is replaced by decentralized, probably reciprocal descriptions. We call this a *mesoscopic model*. It differs from existing microscopic analyses (Beyer and Rudolph, 1997; Altenberg, 1995) in the sense that the considered microscopic details are aggregated such that the linkage of multiple results is possible.

The framework enables the integration of existing theoretical results as well as empirical investigations. Especially the generalization of theoretical results using empirical methods is an encouraging approach to broaden the understanding and to create a foundation for a general comprehension of EA. Overall, the framework aims at a holistic, general model of evolutionary computing.

For a decentralized model certain properties or entities of evolutionary search must be determined that may serve as keystones of evolutionary search processes. In this paper we discuss and analyze diversity as one of those keystones.

**Table 1** General evolutionary algorithm

---

```
1:  $t \leftarrow 0$ 
2:  $P(t) \leftarrow$  create population of size  $\mu$ 
3: evaluate  $P(t)$  using  $f$  and  $dec$ 
4: while termination criterion not fulfilled do
5:    $P' \leftarrow$  select parents for  $\lambda$  offspring from  $P(t)$  using selection operator  $S$ 
6:    $P'' \leftarrow$  create offspring by applying recombination operator  $R$  to  $P'$ 
7:    $P''' \leftarrow$  mutate each individual in  $P''$  by operator  $M$ 
8:   evaluate  $P'''$  using  $f$  and  $dec$ 
9:    $t \leftarrow t + 1$ 
10:   $P(t) \leftarrow$  select  $\mu$  individuals from  $P'''$  (or from  $P''' \circ P(t-1)$ ) using selection operator  $S'$ 
11: end while
12: output best individual in  $P(t)$ 
```

---

## 2 BASIC FRAMEWORK

To make precise statements on the effects and dependencies of general evolutionary algorithms, it is necessary to define a framework for the examined object. This enables the definition of clear properties to describe the connections in the aspired network of understanding.

The primary goal of such a framework is to have a common basis for a broad range of evolutionary algorithms without restrictions to any of the paradigms. The framework proposed here has several predecessors (e.g. Jones, 1995; Bäck, 1996). The general evolutionary algorithm is displayed in Table 1 in pseudo code. The framework is certainly restricted to rather general algorithms and problems—e.g. ruling out multi-objective or time-varying problems and the involved algorithmic techniques.

The following definition describes the general form of an individual.

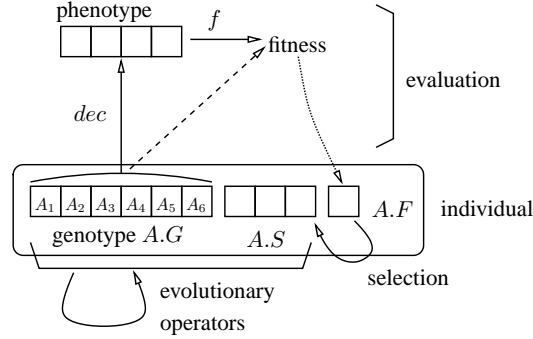
**Definition 1 (Individual)** Each individual is a tuple  $A = \langle G, S, F \rangle$  where  $G \in \mathcal{G}$  is the genotype,  $S \in \mathcal{Z}$  is additional information like strategy parameters, and  $F \in \mathbb{R}$  is the fitness value that was computed last.

For a concrete instance of an algorithm the spaces  $\mathcal{G}$  and  $\mathcal{Z}$  must be defined. Note that  $\mathcal{Z} = \emptyset$  is a valid option. How such an individual interacts with all other components of an evolutionary algorithm is sketched in Figure 1.

The different components of an evolutionary algorithm are formalized in the following three definitions.

**Definition 2 (Operators)** For a genotypical search space  $\mathcal{G}$  and the space  $\mathcal{Z}$  of additional information, a mutation operator is of the form

$$M : \mathcal{G} \times \mathcal{Z} \times \Xi \rightarrow \mathcal{G} \times \mathcal{Z}$$



**Figure 1** Interfaces of an individual to the fitness evaluation, the evolutionary operators, and the selection.

where  $\Xi$  are all possible states of the random number generator. A recombination operator with  $r \geq 2$  parents and  $s \geq 1$  offspring is analogously of the form

$$R : (\mathcal{G} \times \mathcal{Z})^r \times \Xi \rightarrow (\mathcal{G} \times \mathcal{Z})^s.$$

The fitness function cannot evaluate an individual directly but a decoding function must translate the genotype to a phenotypical candidate solution in the space  $\Omega$  on which the fitness function  $f : \Omega \rightarrow \mathbb{R}$  is defined.

**Definition 3 (Decoding function)** A decoding function is of the form

$$dec : \mathcal{G} \rightarrow \Omega$$

with the genotypical search space  $\mathcal{G}$  and the phenotypical search space  $\Omega$ .

Note, that an identity mapping may be chosen for  $dec$  if the phenotype equals the genotype ( $\Omega = \mathcal{G}$ ).

The last components are the selection operators. The following definition takes care of the restriction that selection mechanisms do not introduce new individuals.

**Definition 4 (Selection operator)** A selection operator selecting  $s$  individuals from a population  $P = \langle A^{(1)}, \dots, A^{(r)} \rangle$  has the following form

$$S : (\mathcal{G} \times \mathcal{Z} \times \mathbb{R})^r \times \Xi \rightarrow (\mathcal{G} \times \mathcal{Z} \times \mathbb{R})^s$$

where it holds that  $S(P, \xi) = \langle A^{(index_i)} \rangle_{1 \leq i \leq s}$  with  $1 \leq index_i \leq r$  for  $1 \leq i \leq s$  and  $\xi \in \Xi$ .

Note, that the selection operator is applied to an ordered list of individuals. If such an operator is applied to  $P''' \circ P(t-1)$  in line 10 of the algorithm in Table 1, steady state selection schemes can be realized using a replacement strategy.

A specific algorithm is specified within the framework by the population size  $\mu$ , the number of offspring  $\lambda$ , the number of generations (or evaluations), and the operators with all their parameter settings.

### 3 A KNOWLEDGE REPRESENTATION FOR EVOLUTIONARY PROCESSING

In order to define a knowledge representation for EAs, the first step is a clarification of the relevant objects which knowledge is gathered on. The second step is a discussion and an exact definition of what we actually refer to as “knowledge”. And as a last step the access and the usage of the gathered knowledge needs to be discussed. The first two steps are presented in detail in this section where the surface of the latter step is only scratched and a profound discussion is postponed to a succeeding publication.

The primary purpose of the knowledge representation is an accurate modeling of the characteristics inherent in evolutionary processing—a model of qualitative changes in dependence on parameter settings and problem properties. The characteristics can probably be described best as “natural laws” comparable to the fundamental laws in the natural sciences. Where the objects and entities of physics and chemistry are well known and defined, the relevant entities in the proposed mesoscopic model are not as obvious. Apparently the parameters of the algorithm and the properties of the environment, the optimization problem, are relevant objects. If we want to formulate statements of how good the algorithms perform, any property describing the goals or the performance is of interest, too. Furthermore, in a mesoscopic model any entity summarizing an aspect of the dynamics of the evolutionary processing should also be modeled. We refer to these entities as “*keystones*” or “*cardinal points*” since they are the clue to understand and describe the functioning of EAs in a decentralized manner.

The probably most obvious keystone is the diversity of the population which is examined closer in the second half of this contribution. Other keystones are more vague and the focus of future work. We presume that exploration vs. exploitation, local performance issues, and self-adaptive mechanisms are other relevant keystones. A good starting point for the extraction of new keystones is the analysis of existing theoretical work.

However, both parameters and keystones are only suited to serve as an entity if they have a gradual range of possible values. The mere fact that fitness proportional selection is used does not satisfy this requirement since it does not have a parameter for tuning the selective pressure. Rather, discrete properties are referred to as *switches*. Table 2 provides a few examples of relevant entities and switches. There is a smooth transition between entities and switches and the classification of certain terms may change with a suitable definition of the semantics.

The crucial part of the model is to add “knowledge” to the entities defined in the previous paragraph by bringing certain aspects in relation to other aspects. Since the proposed model is only about qualitative relations the connection of several entities must model how changing the value of at least one entity affects another entity. As an example, increasing the mutation rate can lead under certain circumstances to an increased diversity. In this example we see that the knowledge entry consists of a precondition (including the initial

**Table 2** Examples of entities and switches in the proposed model.

<b>entities (properties with a gradual character):</b>	
<b>parameters:</b>	parental selective pressure, environmental selective pressure, mutation strength (unifying the standard deviation from ES-mutation and the mutation rate from GA-mutation), recombination rate
<b>problem:</b>	HC-hardness (solution quality of a hillclimbing approach)
<b>time:</b>	current fitness level, phenotypical distance to known optima, genotypical distance to known optima
<b>keystones:</b>	diversity, recombinative potential of the population, effective selective pressure, partial convergence, exploitation, exploration, local progress, configuration of schemata/formae in the population, fitness distribution of the population, fitness correlation, self-adaptive potential of the population
<b>goals:</b>	accuracy, required time
<b>switches (properties with a discrete character):</b>	
<b>parameters:</b>	type of environmental selection, type of parental selection, representation and the encoding structure, mutation type, recombination type
<b>problem:</b>	multi-modality, separability

change) and an effect. Additionally, a degree of reliability is assigned to each knowledge entry.

To define the knowledge representation of interdependences in evolutionary algorithms the following definition of a hypergraph is introduced.

**Definition 5 (Hypergraph)** A directed hypergraph is a tuple  $G = (V, E, \gamma)$  where  $V$  is a finite set of vertices,  $E \subseteq \mathcal{P}(V) \times \mathcal{P}(V)$  contains all edges with  $e = (e_1, e_2) \in E$  starting at  $e_1 \neq \emptyset$  and ending at  $e_2 \neq \emptyset$ , and a labeling of the edges  $\gamma : E \rightarrow \Sigma$  with labels  $\Sigma$ .

A directed hypergraph provides the general structure that is necessary to introduce the network of effects and dependencies as a knowledge representation formally.

**Definition 6 (Network of effects)** The network of effects and dependencies is defined as a hypergraph where  $V$  are all entities and all edges  $(e_1, e_2) \in E$  end in one vertex ( $|e_2| = 1$ ). Furthermore edges  $(e_1, e_2)$  are labeled by a tuple of information of the set

$$\Sigma \Big|_{(e_1, e_2)} = \mathit{COND}(e_1) \times \mathit{MOD}(e_2) \times \mathit{CONF}.$$

$\mathit{COND}(e_1)$  denotes all propositional logic formulas using  $\{\wedge, \vee, \neg\}$  and the atoms

$$\begin{aligned} & \{ \mathit{Incr}(v), \mathit{Decr}(v), \mathit{IsHigh}(v), \mathit{IsMedium}(v), \mathit{IsLow}(v) \mid v \in e_1 \text{ is an entity} \} \\ \cup & \{ \mathit{On}(v), \mathit{Off}(v) \mid v \in e_1 \text{ is a switch} \} \end{aligned}$$

where at least one *Incr* or *Decr* occurs. They are the preconditions for the activity (or visibility) of the edge. The effect of the edge on  $e_2$  is described by an element of

$$\mathcal{MOD}(e_2) \in \{Incr(v), Decr(v) \mid v \in e_2 \text{ is an entity}\}.$$

The confidence or the certainty of the represented interdependence is characterized by

$$\mathcal{CONF} = \{\text{proved, high confidence, medium confidence, hypothetical}\}.$$

The current proposition of the network in the definition above uses a rather coarse grained graduation of the entities' values. This could be easily replaced by a finer grading—with several consequences on the effort to gather knowledge and on the inferencing mechanism.

An example for a knowledge entry in the network defined above is the following rule where the precondition is a conjunction of the terms

- (1) *Incr(mutationStrength)*
- (2) *Off(recombinationOperator)* (no recombination)
- (3) *IsMedium(selectivepressure)*

and the effect is described by

$$Incr(diversity).$$

This rather obvious interdependence can also be derived from the experiments in the remaining sections of this work. As a consequence the confidence in the rule is rather high.

Another example for a theoretically derived rule is the following conclusion of the “missing” schema theorem by Altenberg (1995). The rule can be described by the conjunction of preconditions

- (1) *On(probabilisticParentalSelection)*
- (2) *Off(environmentalSelection)*
- (3) *Off(mutationOperator)*
- (4) *IsHigh(recombinativeFitnessCorrelation)*
- (5) *Incr(ratioAboveaverageSchemaPairs)*

where the *recombinativeFitnessCorrelation* denotes the correlation between the fitness of an individual and the average fitness values of the possible parents in the population and the *ratioAboveaverageSchemaPairs* denotes the fraction of schemata (or more general: formae) pairs represented in the population that might be recombined to individuals with above average fitness. The effect of this edge is described by

$$Incr(localProgress).$$

Since this rule is directly derived from a theoretically proved result its confidence is “proved”.

The intended usage of the model is currently rather rudimentary. The primary idea is to initialize the network with certain parameter choices and to assume the changing of

at least one parameter. Then, various rules are triggered either directly by the parameter entities or by the effects of previously triggered rules. In the latter case, it is either possible that a decreasing or increasing effect fulfills a decreasing or increasing precondition or that several decreasing or increasing effects fulfill a precondition “IsLow” or “IsHigh” with a certain probability or confidence. This sequence of triggered rules gives an insight into the possible consequences of the parameter change.

## 4 THE ROLE OF DIVERSITY

In the last section, diversity was proposed as one possible cardinal point or keystone for the processing in evolutionary algorithms within the considered mesoscopic model. This section illustrates the consideration of diversity and sets the stage for the concrete, empirical derivation of “knowledge” concerning the role of diversity in the next section.

Although many publications are concerned with the necessity to keep a high diversity level (for early examples see Goldberg and Richardson, 1987; Cobb, 1990), statements concerning the impact on and effect of diversity are very seldom. However, the entity diversity is presumably crucial for the success or the failure of optimizing with evolutionary algorithms.

Problems occur already when a formal definition of diversity is needed. There are so manifold aspects associated with the term that a single, general definition appears to be impossible. For example, in binary search spaces a standard definition is an entropy measure like the Shannon entropy (Grefenstette, 1987; Kim, 1997; Weicker and Weicker, 2001). This method, however, does only consider each genotypical search space dimension in isolation—the distribution of the individuals within the search space or a correlation between the values of certain search space dimensions are not considered.

Therefore, an alternative diversity measure is used in this work, namely the average pairwise distance of any two distinct individuals in the population,

$$Diversity(P) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n d(A^{(i)}, A^{(j)})$$

for population  $P = \langle A^{(1)}, \dots, A^{(j)} \rangle$ . Since it mediates a better picture of the actual population than an entropy measure we have chosen it for this investigation. It is also a general diversity measure since it can be applied to any search domain where a distance function  $d$  is defined. In this work the Hamming distance is used for binary search spaces and the Euclidean distance in real-valued search spaces. A basically similar measure was introduced by Barker and Martin (2000). A variety of different diversity measures can be found in the work of Herrera and Lozano (1996).

To introduce an entity “diversity” as a keystone of evolutionary processing in the desired network, the interdependence of diversity and the various components of evolutionary algorithms must be clarified. The following considerations are the basis for the empirical investigations in the next section.

First, the effects of diversity are discussed. A rather high diversity is usually associated with a high degree of adaptivity and flexibility during the optimization process. As a



negative extreme the loss of diversity leads to premature convergence and the evolutionary process grinds to a halt. This is due to a very narrow window of the search space that can be reached from the converged population using the evolutionary operators. This leads to the question: How much do the evolutionary operators benefit from a high diversity? Inaccurately, we can argue that the possible offspring of a mutation applied to the individuals of a population depend linearly on the number of individuals. However, a recombinative operator with  $k$  parents and a population size  $n$  has  $\binom{n}{k}$  possible parental combinations; and the number of offspring can be roughly estimated as linearly dependent on this number. As a consequence the recombination suffers more substantially under a low diversity since the possible paths of recombination are chopped off nonlinearly.

Second, the impact of the evolutionary algorithm's components on the diversity is examined. Well understood is the effect of standard selection mechanisms which are always reducing the population's diversity. Probably the takeover time is a good reference point for the change in the diversity (see e.g. Blickle and Thiele, 1997). The mutation operator plays an antagonistic role since it increases the diversity always. For example in dynamic optimization there are techniques to keep a high diversity by a very high mutation rate (Cobb and Grefenstette, 1993). The role of the recombination operator is more complex. From a population geneticist's point of view, a crossover operator has no impact on the diversity of the population since it does not change the frequency of the single alleles (locus-wise diversity). However, this cannot be transferred to the distance based diversity introduced above. Instead, recombination can increase or decrease the diversity. Applied to a well distributed population, uniform crossover and 2-point crossover increase the diversity by introducing recombined individuals. But for a rather contracting recombination operator, like the intermediate recombination (or arithmetic crossover), evolution strategy's genetic repair theory tells us that the operator reduces the effects of bigger mutations considerably (for global intermediate recombination see Beyer, 2001). In the bottom line, this equals a diversity diminishing effect.

There seem to be many facets of diversity worth to be used in a mesoscopic modeling approach. The next section focuses on the empirical investigation of the contracting recombination by examining the hypothesis that there is a threshold where increasing the recombination rate leads to decreasing diversity. In this work we characterize a recombination operator as contracting if a random recombinative walk applied to an arbitrary population converges on one individual with a significant probability. Also, the hypothesis that increasing the mutation rate increases the diversity is an issue in the next section.

## 5 EMPIRICAL KNOWLEDGE DERIVATION

This section demonstrates how knowledge can be derived for the network empirically. Since the idea of the network is the representation of the relevant "natural laws" of evolutionary computation, a good empirical methodology is crucial. Edges in the network are supposed to be true for a wide range of algorithms and problems that meet the respective preconditions. The adjustment of existing theoretical results is straight forward since the assumptions of a lemma or theorem can be transformed directly to preconditions of the knowledge entry. However, for empirical investigations the following rules should be a minimal requirement.

1. Rather distinct algorithms are chosen to allow generalization from the concrete algorithms.
2. Rather distinct benchmark problems are chosen to allow generalization.
3. A concrete hypotheses concerning the influence of entities on other entities must be proposed; the algorithms' parameter combinations must be selected accordingly.
4. A sufficient number of experiments for each algorithm, problem, and parameter combination must be executed and documented.
5. A statistical examination of the data is carried out to approve or disprove the hypothesis. Probably further preconditions can be derived from the empirical investigation concerning the area of validity of the hypothesis.

These requirements for the derivation of knowledge entries are the basis for the examination of the influence of a contracting recombination on the diversity. The experimental setup and the results are discussed in the remainder of this section.

### 5.1 EXPERIMENTAL SETUP

Two algorithms are used to underpin the hypothesis that the diversity decreases for a contracting recombination.

The first algorithm is a (20,100)-evolution strategy (ES) using Gaussian mutation with a fixed standard deviation and a recombination operator applied to create a certain fraction of offspring (in expectation). The values for the standard deviation are

$$\delta, \frac{1}{2}\delta, \frac{1}{4}\delta$$

where  $\delta = \frac{\text{range}}{2\sqrt{\text{dimension}}}$  (as proposed by Schwefel (1977) as initial standard deviation—however Bäck (1996) proposes smaller values), *range* is the distance of the search space boundaries in one dimension, and *dimension* is the dimensionality of the search space. As recombination operators, 2-point crossover, uniform crossover, and the contracting intermediate recombination are used. The percentage of individuals created by recombination is an element of the set {0.0,0.3,0.7}. Adaptation and self-adaptation techniques have been omitted to guarantee the independence of the mutation step-width.

The second algorithm is a genetic algorithm (GA) with population size 100, parental 3-ary tournament selection, and standard binary encoding using 16 bits per search space dimension. The values for the mutation rate are

$$\frac{1}{2l}, \frac{1}{l}, \frac{2}{l}$$

where  $l = 16 \cdot \text{dimension}$ . As recombination operator 2-point crossover, uniform crossover, and a biased version of the uniform crossover are chosen. In the latter biased crossover 2.5% of the bits are chosen from the worse parent and the other bits from the better parent. Again, the percentage of individuals created by recombination is an element of the set {0.0,0.3,0.7}.

The considered problems are the sphere, the weighted sphere, Schwefel's double sum, and the Rosenbrock function with *dimension* = 30. All four functions are unimodal and

continual in the phenotypic search space  $\Omega$ . The first two functions are separable, the latter two non-separable.

Note, that the neighborhood structure introduced by the GA mutation results in a multimodal landscape on genotype space  $\mathcal{G}$ . The different neighborhood structures are used intentionally since we are interested in general “natural” laws. However, the impact of the genotypic neighborhood is omitted here and will be considered more thoroughly in future examinations.

The comparison of diversity values in experiments using differently adjusted optimizers is a difficult task since e.g. the change of the standard deviation also changes the speed of the optimization process. Therefore, in this investigation the diversity values are compared relatively to the fitness level reached. As a consequence, for each problem a number of “milestones” are defined, i.e. relevant fitness values at which the diversity values of two algorithms are compared.

Each algorithm (or setup of the algorithm) is run 30 times with different initial seeds for the random number generator. A statistical t-test compares those diversity values at the generations where the respective milestone is reached first. An error probability of less than 0.05 is considered to be significant. In the investigation the t-test was only applied to milestones that have been reached by all 30 experiments.

## 5.2 EXPERIMENTAL RESULTS

The results for the ES variant are shown in Figures 2, 3, 4, and 5. The results for the GA are shown in Figures 6, 7, 8, and 9. In all figures the results of the t-tests are shown with an indication concerning the validity area of the hypothesis. The validity areas are displayed in the lower right corner of each figure.

Concerning the intuitive hypothesis in Section 3, that increasing the mutation strength increases the diversity (without recombination), is supported by the data of GA and ES. However, for the ES and the weighted sphere there is a restriction for big step sizes and a close to optimum performance. As a consequence the validity range of the step size could be constrained to medium and low values.

The results concerning the hypothesis on the contracting recombination are not as clear as the results above. In case of the ES, the predicted effect occurs for all problems, between recombination rate 0.3 and 0.7, and independent of the mutation strength. This is significantly different from the behavior of the non-contracting uniform and 2-point crossover operators. However, the behavior of the contracting biased uniform crossover in the GA is different from the ES case. In the GA experiments, the increase of the diversity with increasing recombination rate appears to be much more severe for the 2-point and uniform crossover than it is the case for the ES experiments. The biased mutation alleviates this effect such that there is often only a reduced or no significance for the increase. In case of higher mutation strength, the desired effect of decreasing diversity takes even place.

Apparently, the conversion of the increasing diversity is due to different effects in the two considered evolutionary algorithms. The ES recombination shows the predicted behavior. But the biased uniform crossover has a different structure creating a contracting effect. Where the intermediate recombination stays inside an area determined by the individuals

in the population, the biased crossover only reduces the number of outliers. Therefore, it appears that the recombination is not able to converse the increasing diversity without the help of a sufficiently big mutation strength.

By considering the intersection of both effects, a hypotheses entry into the network would contain as precondition part the conjunction of following predicates:

- (1)  $On(unimodal)$
- (2)  $IsMedium(envSelectivePressure) \vee IsMedium(parentSelectivePressure)$
- (3)  $On(contractingRecombination)$
- (4)  $IsHigh(closeToOptimumFitness)$
- (5)  $IsMedium(mutationStrength) \vee IsHigh(mutationStrength)$
- (6)  $Incr(recombinationRate)$

The effect is described by the following predicate:

$$Decr(diversity).$$

For a much clearer description of the underlying “natural law” it is necessary to carry out further experiments. These experiments should consider the comparability of the two contracting operators with regard to the search space structure and the dynamics of the contraction as well as the influence of the different selective pressure (and the different selection schemes). Also the direct connection of parameters and diversity should be questioned: The identification of a sensible granular definition for the entities “exploration” and “exploitation” could lead to an indirect impact of parameters onto the diversity.

## 6 CONCLUSION AND FUTURE WORK

This article introduces qualitative modelling to research in evolutionary computation. It proposes a network as knowledge representation for empirical and theoretical results concerning the local effects and dependencies in evolutionary algorithms. Where global models seem to have severe limitations, we believe that this decentralized gathering of results will lead to a more general understanding of evolutionary algorithms. Especially the interweaving of results from different backgrounds (e.g. schema theory vs. ES theory) and different granularity concerning their force of expression (e.g. population based results vs. individual based results) is the first approach toward an integrating, general theory. Future work will show whether this leads to a networked understanding of evolutionary computing.

The exemplary investigation of the diversity as an indirect property of an evolutionary search has shown how the framework may be filled. Also it can be seen that the generalization of theoretical results on a broad empirical basis is actually possible. However, also the problems of the presented approach are obvious: Experimental design and evaluation are very tedious tasks and the missing profound basis for the definition of keystones requires a phase of trial and error. At the current state, a useful model can only be constructed by an iterative refinement process.

As a consequence, this work is only a first proposal for an integrating qualitative model. Hopefully it leads to a fruitful discussion concerning standards of empirical evaluation and an overall model. The definition and the form of the network needs to be revised on the basis of more researchers' input.

The utility of the model will be a qualitative reflection of the underlying principles of evolutionary processing and the identification of probably still unknown connections and fundamental laws. Also it could serve as a tool to support the failure analysis of insufficient evolutionary algorithms.

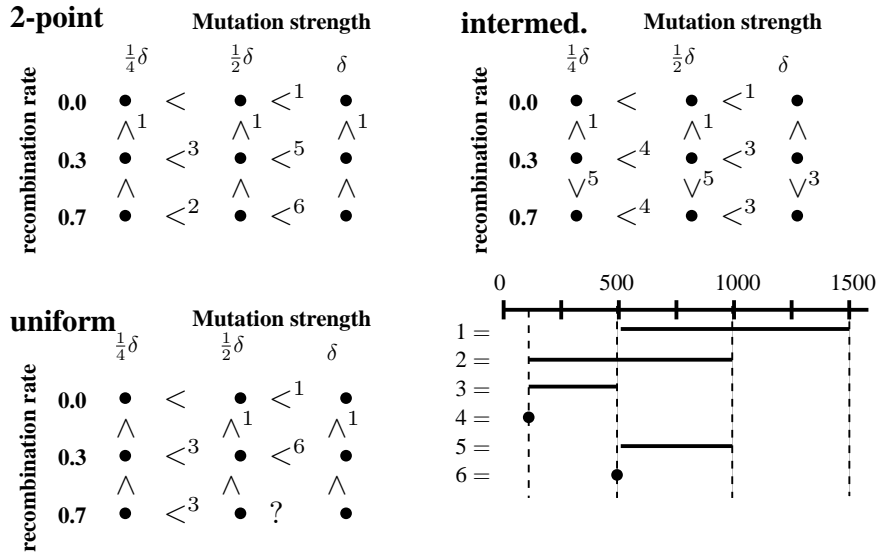
The major part of the article is concerned with a rather big and unifying model of interactions. The possible impact of such a model on EA design is an open question. Currently we plan to use the "big" model for educational purposes. To support design of EAs, more focused small models of interactions could be necessary—focusing on a special EA paradigm and/or on a special area of application.

Future work has to deal with the investigation of more keystones—first candidates are exploitation, exploration, role of schemata/formae, and certain aspects of adaptation and self-adaptation. Also the inference mechanism for the network needs to be developed. If the qualitative model matures in the long run, it should be possible to transfer the gathered knowledge into a quantitative model.

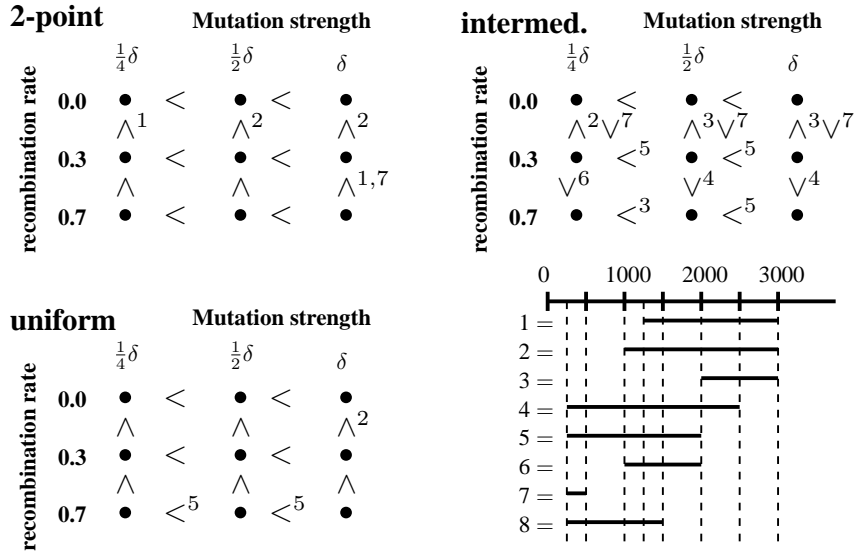
Besides the introduction of a qualitative model, this paper is a call to the research community to put our experimental investigations on a better basis. If empirical results shall serve as a second foundation of evolutionary computing—next to theory as a first foundation—we must stop the simple reporting of what we have done and what we have seen in favor of a solid empirical investigation and interpretation of our data. The proposed formalism is a necessary tool to archive and to connect the results.

## References

- Altenberg, L. (1995). The schema theorem and Price's theorem. In L. D. Whitley and M. D. Vose (Eds.), *Foundations of Genetic Algorithms III*, San Francisco, CA, pp. 23–49. Morgan Kaufmann.
- Bäck, T. (1996). *Evolutionary Algorithms in Theory and Practice*. New York: Oxford University Press.
- Barker, A. L. and W. N. Martin (2000). Dynamics of a distance-based population diversity measure. In *Proc. of the 2000 Congress on Evolutionary Computation*, Piscataway, NJ, pp. 1002–1009. IEEE Service Center.
- Beyer, H.-G. (2001). *The theory of Evolution Strategies*. Berlin: Springer.
- Beyer, H.-G. and G. Rudolph (1997). Local performance measures. In T. Bäck, D. B. Fogel, and Z. Michalewicz (Eds.), *Handbook of Evolutionary Computation*, pp. B2.4:1–27. Bristol, New York: Institute of Physics Publishing and Oxford University Press.
- Blickle, T. and L. Thiele (1997). A comparison of selection schemes used in evolutionary algorithms. *Evolutionary Computation* 4(4), 361–394.
- Cobb, H. G. (1990). An investigation into the use of hypermutation as an adaptive operator in genetic algorithms having continuous, time-dependent nonstationary environments.



**Figure 2** The diversity behavior of the ES for the Rosenbrock function. The points mark combinations of mutation strength and recombination rate. The inequality symbols show where a significantly higher diversity could be detected. The small numbers indicate the range where the inequality is valid. The question mark labels connections where no significant difference in the diversity can be verified.



**Figure 3** The diversity behavior of the ES for the Sphere function.

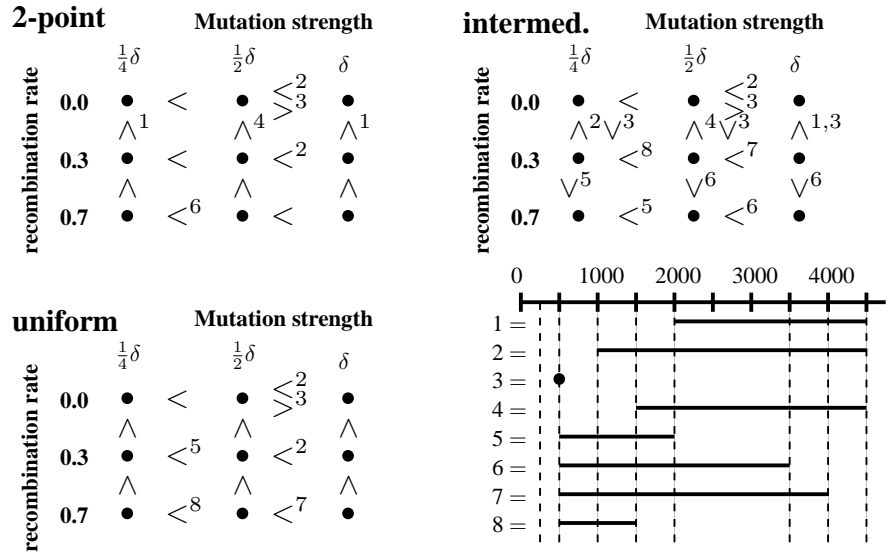


Figure 4 The diversity behavior of the ES for the Weighted Sphere function.

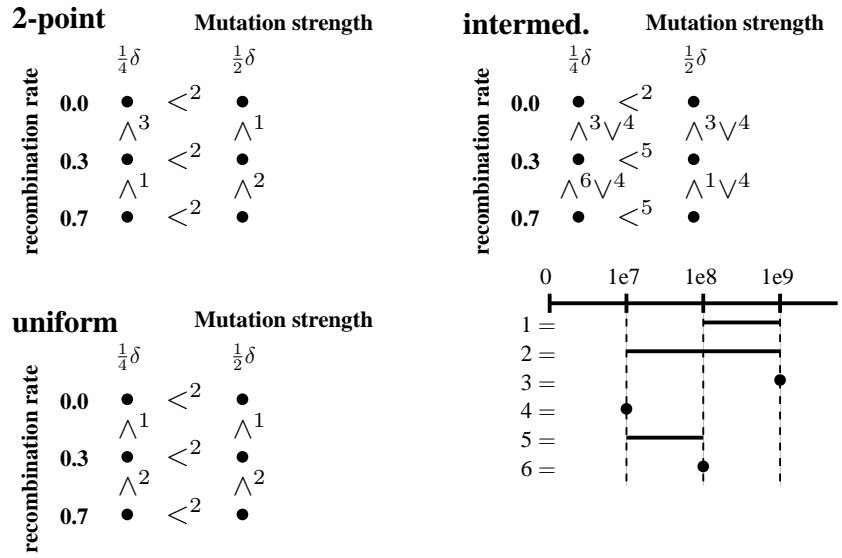


Figure 5 The diversity behavior of the ES for the Schwefel double sum function.

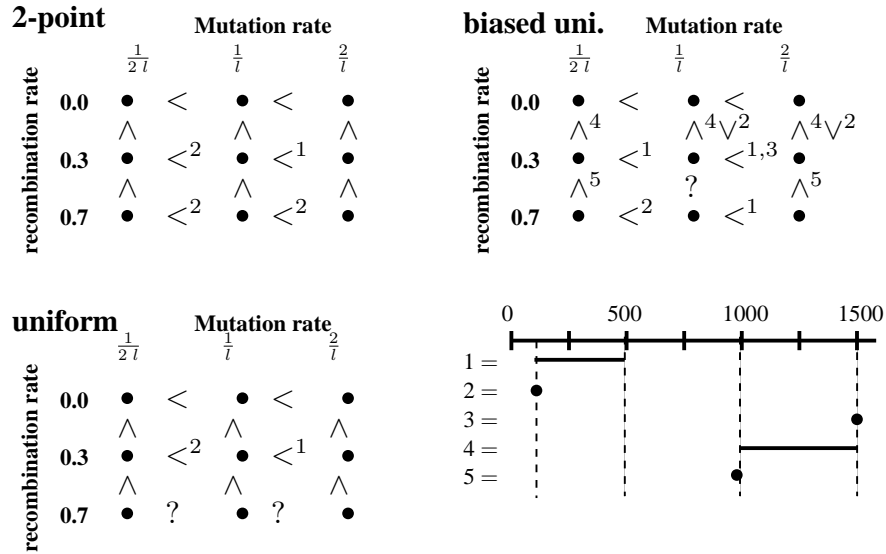


Figure 6 The diversity behavior of the GA for the Rosenbrock function.

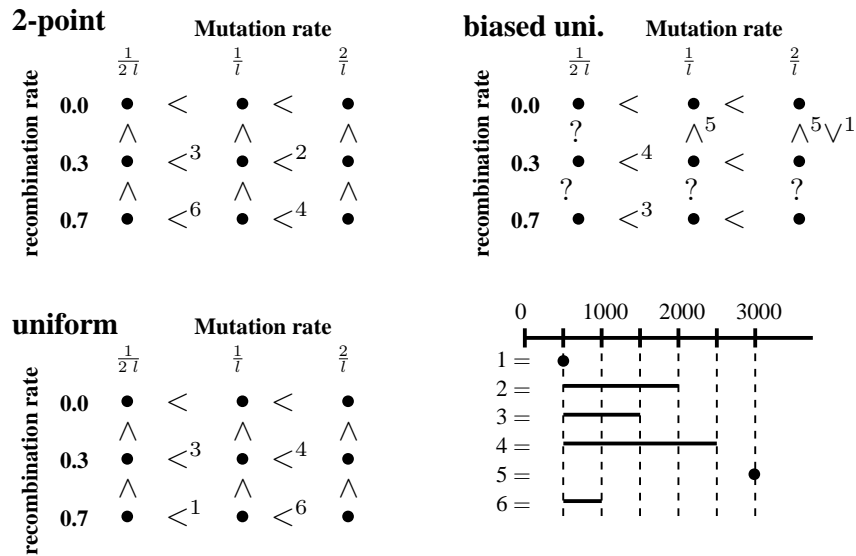


Figure 7 The diversity behavior of the GA for the Sphere function.



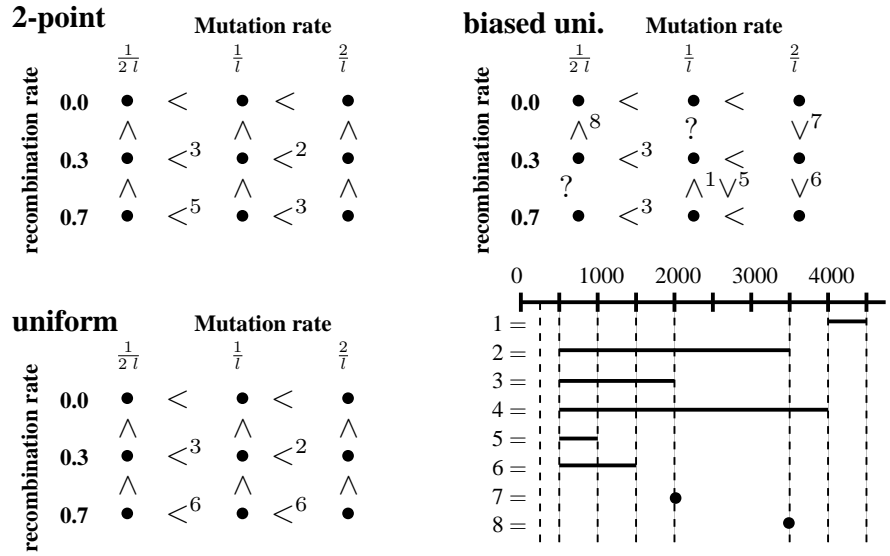


Figure 8 The diversity behavior of the GA for the Weighted Sphere function.

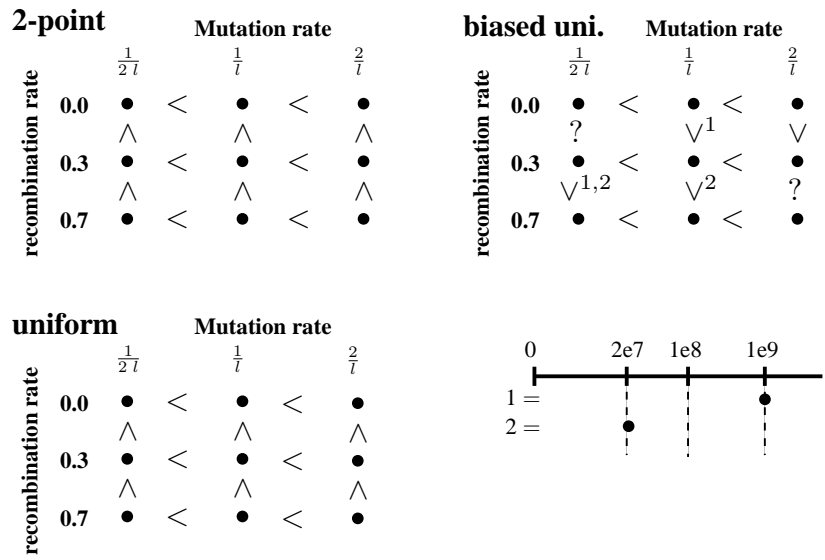


Figure 9 The diversity behavior of the GA for the Schwefel double sum function.

- Technical Report 6760 (NLR Memorandum), Navy Center for Applied Research in Artificial Intelligence, Washington, D.C.
- Cobb, H. G. and J. J. Grefenstette (1993). Genetic algorithms for tracking changing environments. In S. Forrest (Ed.), *Proc. of the Fifth Int. Conf. on Genetic Algorithms*, San Mateo, CA, pp. 523–530. Morgan Kaufmann.
- Goldberg, D. E. and J. Richardson (1987). Genetic algorithms with sharing for multimodal function optimization. In J. J. Grefenstette (Ed.), *Genetic algorithms and their applications: Proc. of the Second Int. Conf. on Genetic Algorithms*, Hillsdale, NJ, pp. 41–49. Lawrence Erlbaum Assoc.
- Grefenstette, J. J. (1987). Incorporating problem specific knowledge into genetic algorithms. In L. Davis (Ed.), *Genetic Algorithms and Simulated Annealing*, pp. 42–60. Los Altos: Morgan Kaufmann.
- Herrera, F. and M. Lozano (1996). Adaptation of genetic algorithm parameters based on fuzzy logic controllers. In F. Herrera and J. L. Verdegay (Eds.), *Genetic Algorithms and Soft Computing*, pp. 95–125. Heidelberg: Physica-Verlag.
- Jones, T. (1995). *Evolutionary Algorithms, Fitness Landscapes and Search*. Ph. D. thesis, The University of New Mexico, Albuquerque, NM.
- Kim, J. T. (1997). Distance distribution complexity: A measure for the structured diversity of evolving populations. In C. G. Langton and K. Shimohara (Eds.), *Artificial Life V: Proc. of the Fifth Int. Workshop on the Synthesis and Simulation of Living Systems*, Cambridge, MA, pp. 281–288. The MIT Press.
- Poli, R. (1999). Schema theorems without expectations. In W. Banzhaf, J. Daida, A. E. Eiben, M. H. Garzon, V. Honavar, M. Jakiela, and R. E. Smith (Eds.), *Proc. of the Genetic and Evolutionary Computation Conf. GECCO-99*, San Francisco, CA, pp. 806. Morgan Kaufmann.
- Radcliffe, N. J. and P. D. Surry (1995). Fitness variance of formae and performance prediction. In L. D. Whitley and M. D. Vose (Eds.), *Foundations of Genetic Algorithms III*, San Mateo, CA, pp. 51–72. Morgan Kaufmann.
- Rajagopalan, R. (1986). Qualitative modeling and simulation: A survey. In E. J. H. Kerckhoffs (Ed.), *Proc. of the European Conference "AI Applied to Simulation"*, San Diego, CA, pp. 9–26. Simulation Councils.
- Rudolph, G. (1998). Finite Markov chain results in evolutionary computation: A tour d’horizon. *Fundamenta Informaticae* 35, 67–89.
- Schwefel, H.-P. (1977). *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie*. Basel, Stuttgart: Birkhäuser. Volume 26 of Interdisciplinary Systems Research, German.
- Shapiro, J. L. (2001). Statistical mechanics theory of genetic algorithms. In L. Kallel, B. Naudts, and A. Rogers (Eds.), *Theoretical Aspects of Evolutionary Computing*, pp. 87–108. Berlin: Springer.
- Weicker, K. and N. Weicker (2001). Burden and benefits of redundancy. In W. N. Martin and W. M. Spears (Eds.), *Foundations of Genetic Algorithms 6*, pp. 313–333. San Francisco: Morgan Kaufmann.