UNIVERSITY OF DORTMUND

REIHE COMPUTATIONAL INTELLIGENCE

COLLABORATIVE RESEARCH CENTER 531

Design and Management of Complex Technical Processes and Systems by means of Computational Intelligence Methods

A Unified Model of Non-Panmictic Population Structures in Evolutionary Algorithms

Joachim Sprave

No. CI-55/99

Technical Report ISSN 1433-3325 January 1999

Secretary of the SFB 531 · University of Dortmund · Dept. of Computer Science/XI
44221 Dortmund · Germany

This work is a product of the Collaborative Research Center 531, "Computational Intelligence", at the University of Dortmund and was printed with financial support of the Deutsche Forschungsgemeinschaft.

1 Motivation

In most variants of Evolutionary Algorithms (EAs), multiple search points are considered each iteration. In analogy to nature, the search points are called a population of (artificial) individuals. Due to the fact that computers were slow and mostly sequential machines when these algorithms were developed, the inherent parallelism of populations had to be serialized in the first implementations. Later on, when parallel computers became affordable for universities, researchers began to parallelize EAs again. Since the terminology of the sequential algorithms had more or less settled at that time, the parallel versions were described by an extension of the sequential terminology where a generalization would have been appropriate instead. Surprisingly, this generalization has not been done ever since, although there were approaches to a common terminology of parallel EAs. The work described here is an attempt to provide a unified model of population structures, independently of their parallelization properties. This paper starts with a brief overview of other approaches to model, classify, or analyze population structures. Then, a general framework for the formal description of population structures is introduced, followed by examples of common population structures expressed by means of the given model. Finally, as an application of the formal framework, a method for calculating growth rates and takeover times of arbitrary population structures is presented.

2 Parallelism and Population Structures

Traditional EAs allow each individual in a population to interact (compete or mate) with any other. In biological terms, a population like this is called *panmictic*. This fully connected interaction scheme is not very well suited for a parallel implementation.

On the other hand, parallelizing a sequential algorithm usually means that the parallel version, given the same input as the original algorithm, produces the same output, as well. Since every parallel algorithm can be run time sliced on a sequential machine, the maximum speed-up is limited linearly by the number of processors. In other words, parallelization does not change the quality of results, it just delivers the same results in less time.

Except for master/slave parallelization of fitness function evaluations, all parallel EA approaches violate the same-input-same-output rule. The panmictic approach is replaced by structured populations which can be evaluated in parallel with a reduced need for communication.

First attempts to a unified terminology of non-panmictic population structures were made by Gorges-Schleuter [1], who presented a rough classification of parallel EAs. She identified three major models of parallelization in EAs:

- Island Model:
  The population consists of separated subpopulations. Each subpopulation is a panmictic EA. A limited amount of genetic information is exchanged between arbitrary subpopulations.
- Stepping Stone Model:
  The population consists of separated, panmictic subpopulations. A non-total neighborhood relation is defined on the subpopulations. A limited amount of genetic information is exchanged between adjacent subpopulations.
- Neighborhood Model:
  A non-total neighborhood relation is defined on the individuals. Individuals interact (mate, compete) with adjacent individuals only.

Today, most authors do not discriminate between the island and stepping stone model. Instead, either the term island model or migration model is used for an EA with panmictic subpopulations. The neighborhood model is sometimes also called diffusion model.

In the past, work on population structures always emphasized a particular model. For panmictic subpopulations, there are mostly empirical studies, e.g. [2, 3, 4, 5]. In a theoretical approach, Cantú-Paz [6] presented optimal subpopulation sizes for some special instances of Genetic Algorithms.

Neighborhood models were analyzed with respect to local selection schemes [7, 8] as well as neighborhood shapes in grid topologies [9, 10].

3 A Hypergraph-Based Model of Population Structures

3.1 Hypergraphs

The definitions in this section base on the theory of hypergraphs as defined by Berge [11] in the 60s.

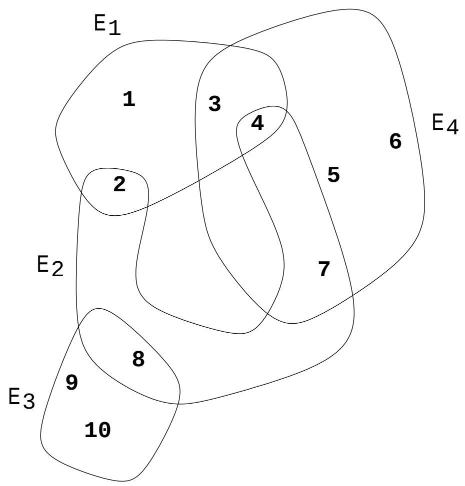

Figure 1: A hypergraph with vertices $X = \{1, \ldots, 10\}$ and edges $E_1 = \{1, 2, 3, 4\}$, $E_2 = \{2, 4, 7, 8\}$, $E_3 = \{8, 9, 10\}$, and $E_4 = \{3, 4, 5, 6, 7\}$.

Graphs are usually defined as sets of pairs of a base set. Elements of the base set are called *vertices*, and pairs of vertices are called *edges*. The basic idea of hypergraphs is the extensions of edges from pairs to arbitrary subsets of the set of vertices:

Def. 1 Let be $X = \{x_1, x_2, \ldots, x_n\}$ a finite set, and $\mathcal{E} = (E_i \,|\, i \in I)$ a family of subsets of X. The family $\mathcal{E}$ is called *hypergraph on* X, if

$$E_i \neq \emptyset \; \forall i \in I \tag{1}$$

$$\bigcup_{i \in I} E_i = X \; . \tag{2}$$

The pair $H = (X, \mathcal{E})$ is called hypergraph, and $|X|$ is the *order* of H.

Two vertices x_k and x_l of an hypergraph are *adjacent*, if they are contained in common edge, formally

$$\exists i : \quad x_k \in E_i \wedge x_l \in E_i \; . \tag{3}$$

For each hypergraph $H = (X, \mathcal{E})$, there exists a *dual Hypergraph* $H^* = (E, \mathcal{X})$, with vertices $e_1, \ldots, e_m$ and edges $\{X_1, \ldots, X_n\}$, given that:

$$\forall j \in \{1, \ldots, n\} : \quad X_j = \{e_i | i \leq m, E_i \ni x_j\} \; , \tag{4}$$

where the vertices e_i correspond with the edges E_i of H, and the edges X_j with the vertices x_j of H, respectively.

3.2 Representations of Hypergraphs

The efficiency of algorithms on simple graphs is usually based on a suitable representation of a given graph. Many algorithms are defined on adjacence lists or adjacence matrices. These representations cannot be easily transfered to hypergraphs, because the information which edge two vertices are connected by is lost.

A loss-free representation of a hypergraph is its incidence matrix. The columns of the matrix represent the edges of the hypergraph and the lines the vertices. If H is a hypergraph with edges m and vertices n, then it is described by the matrix A with

$$a_{i,j} = \begin{cases} 1 & \text{if} \quad x_i \in E_j \\ 0 & \text{else.} \end{cases} \tag{5}$$

The indicidence matrix of the hypergraph given in Fig. 1 is

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The adjacence of two vertices x_i and x_k can be calculated as scalar product of the ith that kth row:

$$< a_{i,\cdot}, a_{k,\cdot} >= 1 \Leftrightarrow \exists E_j : x_i \in E_j \wedge x_k \in E_j. \tag{6}$$

The vertex adjacence matrix B with $b_{i,k} =< a_{i,\cdot}, a_{k,\cdot} >$ can be obtained as follows:

$$B = AA^\top \tag{7}$$

Accordingly one can express the edge adjacence matrix as scalar product of the appropriate lines of the incidence matrix. One receives the matrix as

$$B^* = A^\top A. \tag{8}$$

B^* is as well the vertex adjacence matrix of the dual hypergraph H^*. The matrix B^*, for example, allows the calculation of the diameter with

$$\delta(H) = min\{k | \sum_{i=0}^{k} {B^*}^i = K_n\} \tag{9}$$

where K_n is the adjacence matrix of the complete (simple) graph with n vertices.

3.3 Modeling of Population Structures by Means of Neighborhood Graphs

Non-panmictic population models, as published in the past, usually base on regular geometric structures. Frequently, rings, tori, or cubes are used to place the individuals or subpopulations on. Since all these structures can be described as meshes, there is already a natural notion of adjacency. For neighborhood models, the most popular population layout is a two-dimensional grid folded to a torus.

The neighborhood relation in torus shaped populations is not necessarily limited to adjoining mesh point, but it is more often defined by means of a distance measure in a two-dimensional plane. Clearly, once the neighborhood relation is given, it can be represented by a graph. While a graph representation seems to be appropriate for neighborhood models, it does not very well reflect the structure of coarse grained models. A neighborhood graph of a migration model contains large cliques, i.e. fully connected subgraphs. A formal definition of populations structures by means of graphs is omitted here in favor to the following, more general model which is based on the theory of hypergraphs.

4 Modeling of Population Structures by Means of Hypergraphs

4.1 Definitions

Before we give a definition of population structures, we must state what a population is in this context:

Def. 2 Let be A the search space of a given objective function, S the space of additional state information needed by the optimization algorithm under consideration. Then an element of the family $A \times S$ is called an *individual.* A family of individuals $(p_0^t, \ldots, p_{\lambda-1}^t), p_i^t \in A \times S$ is called *population at time* t or *population at generation* t.

The state information may contain, e.g., the mutation variances of an Evolution Strategy.

Since the actual values of individuals are not needed for the definition of population structures, it is sufficient to identify individuals by their indices in the population.

Def. 3 A *population structure* Π on a population P with $|P| = \lambda$ is a triple $(X, \mathcal{E}, \mathcal{Q})$, consisting of a hypergraph $(X, \mathcal{E}), X = \{0, \ldots, \lambda - 1\}, \mathcal{E} \subseteq \mathcal{P}(X)$, and a partition $\mathcal{Q} \subset \mathcal{P}(\mathcal{X})$ of X with $|\mathcal{Q}| = |\mathcal{E}|$. The hyperedge E_i is called *deme* of the elements from $Q_i \in \mathcal{Q}$.

As the term *deme* indicates, each hyperedge contains the potential parents of the individuals in the corresponding element of the partition. The duality of hypergraph is also reflected by the definition of population structures. Since the partition $\mathcal{Q}$ can be also interpreted as a hypergraph, there are two matrices associated with population structures:

- $\overline{E} \in \mathbb{B}^{(\lambda \times m)}$: Incidence matrix of the hypergraph $(X, \mathcal{E})$.
- $\overline{Q} \in \mathbb{B}^{(\lambda \times m)}$: Incidence matrix of the hypergraph $(X, \mathcal{Q})$.

Consider the relation *i is a potential parent of j*. The adjacence matrix of the associated simple graph can be written as

$$A \quad = \quad \overline{E}\,\overline{Q}^{\top} \ \in \mathbb{B}^{(\lambda \times \lambda)}, \quad \lambda = |X| \tag{10}$$

. Lets calls this matrix the individual adjacence matrix. The dual matrix, the deme adjacence matrix, describes the relation *Q_i has potential offspring in Q_j*. It can be calculated as

$$A^* \quad = \quad \overline{Q}^{\top}\overline{E} \ \in \mathbb{B}^{(m \times m)}, \quad m = |\mathcal{Q}| \tag{11}$$

As can be seen, this modeling approach adjusts to the scale of a particular population structure.

Def. 4 A *path from i to j* in a population structure $\Pi = (X, \mathcal{E}, \mathcal{Q})$ is a finite sequence of 2-tuples $(i_\nu, j_\nu), \nu = 0, \ldots, l$ with $i_0 = i, j_l = j$, and $i_\nu \in Q_k \Rightarrow j_{\nu-1} \in E_k, 1 \leq \nu \leq l$. The value l is the length of the path (i_ν, j_ν).

If there exists a path from i to j, there is at least one shortest path. The longest shortest path between two individuals is called the *diameter* of the population structure. Obviously, the diameter is as well the elitist takeover time of a population structure. The diameter can be calculated as the diameter of the simple graph G with $(x, y) \in G$ if $\exists \nu \in \lambda\mathbb{Z} : x \in E_\nu \cap y \in Q_\nu$. Let be $\overline{E}$ the incidence matrix of $H = (X, \mathcal{E})$, $\overline{Q}$ the incidence matrix of $J = (X, \mathcal{Q})$. Since $g_{i,j} = \sum_{i=0}^{N-1} \overline{e}_{x,i} \cdot \overline{q}_{i,y} =< h_{x,\cdot}, q_{\cdot,y} >$, the adjacence matrix of G is $\overline{E}\,\overline{Q}^{\top}$.

5 Hypergraph Models of Population Structures

5.1 Panmixis

The most common population structure in Evolutionary Algorithms is panmictic:

$$\begin{aligned} \Pi_{\text{pan}} &= (X, \mathcal{E}, \mathcal{Q}) \\ X &= (0, \ldots, \lambda - 1) \\ \mathcal{Q} &= (X) \\ \mathcal{E} &= (X) \end{aligned} \tag{12}$$

5.2 Migration and Pollination

In the following, the population structure of a migration model without isolation is presented. Let be $P = \{0, \ldots, \lambda - 1\}, \lambda \in \mathbb{N}$ a population, $r \in \mathbb{N}$ the number of subpopulations with $\lambda = r\nu$, $r, \nu \in \mathbb{N}$, and $m \in \{1, \ldots, \nu\}$ the number of migrants between adjacent subpopulations. The subpopulations are $Q_i = \{i\nu, \ldots, i\nu + \nu - 1\}, i = 0, \ldots, r - 1$.

Let be $M_{s \to t}$ the set of migrants from subpopulation Q_s to subpopulation Q_t. If $M_{s \to t} \neq \emptyset$, there is a migration path from Q_s to Q_t.

A proper model of migration does not keep the migrants in their source subpopulation after migration, i.e. the migrants are not in the set of potential parents of their successors. This is modeled by subtracting the set of migrants from their source populations hyperedge. The migration model can now be described as follows:

$$\begin{aligned} \Pi_{\text{migr}} &= (X, \mathcal{E}, \mathcal{Q}) \\ X &= (0, \ldots, \lambda - 1) \\ Q_i &= \{i\nu, \ldots, i\nu + \nu - 1\} \\ \mathcal{E} &= (E_0, \ldots, E_{r-1}) \\ E_i &= Q_i \,\cup\, \textstyle\bigcup_{s=0}^{r-1} M_{s \to i} \,\setminus\, \bigcup_{t=0}^{r-1} M_{i \to t}, \\ & \quad M_{s \to i} \subset Q_s,\ M_{i \to t} \subset Q_i \end{aligned} \tag{13}$$

If the subtraction of the migrants is omitted, one obtains a pollination model:

$$\begin{aligned} \Pi_{\text{poll}} &= (X, \mathcal{E}, \mathcal{Q}) \\ X &= (0, \ldots, \lambda - 1) \\ Q_i &= \{i\nu, \ldots, i\nu + \nu - 1\} \\ \mathcal{E} &= (E_0, \ldots, E_{r-1}) \\ E_i &= Q_i \,\cup\, \textstyle\bigcup_{s=0}^{r-1} M_{s \to i}, \quad M_{s \to i} \subset Q_s \end{aligned} \tag{14}$$

The difference to the model above is that the migrants are also potential parents of the next generation of their source subpopulation. Since genetic information is copied instead of moved, the biological analogy of plants spreading pollen is more accurate than that of migrating animals. A pollination model is sketched in Fig. 2, with

$$\begin{aligned} X &= \{0, \ldots, 15\} \\ \mathcal{Q} &= \{\{0, \ldots, 3\}, \{4, \ldots, 7\}, \{8, \ldots, 11\}, \{12, \ldots, 15\}\} \\ \mathcal{E} &= \{\{0, \ldots, 3, 5, 14\}, \{4, \ldots, 7, 3, 10\}, \{8, \ldots, 11, 4, 15\}, \{12, \ldots, 15, 0, 9\}\} \end{aligned} \tag{15}$$

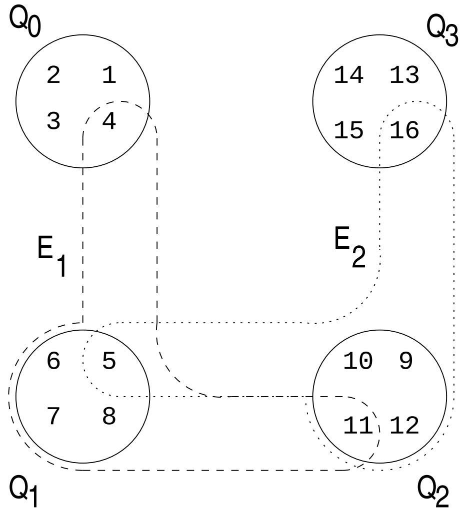

Figure 2: Sketch of a pollination population structure with four subpopulations in a ring. For the sake of clarity, the hyperedges E_0 and E_3 are omitted.

5.3 Fine Grained Models

The majority of the fine grained approaches published in literature uses grid or torus topologies. The neighborhood relation is derived from a topological measure. A simple yet widely used topology, a one-dimensional ring with λ individuals and a neighborhood radius of ρ, is described by the following population structure:

$$\begin{array}{lcl} \Pi_{\text{ring}} & = & (X, \mathcal{E}, \mathcal{Q}) \\ X & = & (0, \ldots, \lambda - 1) \\ \mathcal{Q} & = & (\{x_0\}, \ldots, \{x_{\lambda-1}\}) \\ \mathcal{E} & = & (E_0, \ldots, E_{\lambda-1}) \\ E_i & = & (i \overset{\lambda}{-} \rho, \ldots, i \overset{\lambda}{+} \rho) \end{array} \tag{16}$$

where $\overset{\lambda}{-}$ and $\overset{\lambda}{+}$ are calculated in the factor group $\lambda\mathbb{Z}$. Examples for symmetrical contiguous neighborhoods on two-dimensional meshes are given in [9] and [10].

The mapping of populations in the two-dimensional Euclidean plane usually results from biological analogies[1], in some cases indirectly by the notion of cellular automata[12]. Actually, grid induced neighborhoods are just special cases of graph induced neighborhoods. For a given maximum distance ρ, any connected graph $G \subseteq X \times X$ induces the following population structure:

$$\begin{array}{lcl} \Pi_{\text{graph}} & = & (X, \mathcal{E}, \mathcal{Q}) \\ X & = & (0, \ldots, \lambda - 1) \\ \mathcal{Q} & = & (\{x_0\}, \ldots, \{x_{\lambda-1}\}) \\ \mathcal{E} & = & (E_0, \ldots, E_{\lambda-1}) \\ E_i & = & \{j \mid d_G(e_i, e_j) \le \rho\}, \quad 1 \le \rho \le \lambda/2 \end{array} \tag{17}$$

where $d_G(x, y)$ is the length of the shortest path from x to y in G.

A simple neighborhood model is sketched in Fig. 3, with

$$\begin{aligned} X &= \{0,\ldots,15\} \\ \mathcal{Q} &= \{\{0\},\{1\},\{2\},\{3\},\{4\},\{5\},\{6\},\{7\}\} \\ \mathcal{E} &= \{\{7,0,1\},\{0,1,2\},\{1,2,3\},\{2,3,4\},\{3,4,5\},\{4,5,6\},\{5,6,7\},\{6,7,0\}\} \end{aligned} \tag{18}$$

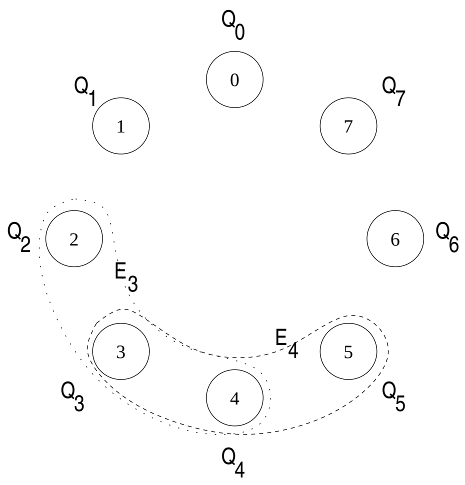

Figure 3: Sketch of a neighborhood population structure with eight individuals in a ring. For the sake of clarity, only the hyperedges E_3 and E_4 are drawn.

5.4 Fine Grained Models

6 Modeling Isolation Times

The definition of population structures given above does not allow the modeling of isolation times. Using edge weights as in classical flow problems results in a cumulation over the time, averaging out some of the effects of isolation. Therefore, the given definition is extended not only to allow isolation, but any change of the population structure over time. This is achieved by replacing the hypergraph H by a sequence of hypergraphs $H^t, t \in \mathbb{N}$, where $(X, \mathcal{E}^t, \mathcal{Q})$ is the population structure applied to generate generation number $t+1$ from generation number t.

Def. 5 A *dynamic population structure* Π on a population P with $|P| = \lambda$ is a triple $(X, \mathcal{E}^t, \mathcal{Q})$, consisting of a sequence of hypergraphs $H^t = (X, \mathcal{E}^t)$, $X = \{0, \ldots, \lambda - 1\}$, $\mathcal{E}_t \subseteq \mathcal{P}(X)$ and $\forall t \in \mathbb{N} : |\mathcal{E}^t| = m$, as well as a partition $\mathcal{Q} \subset \mathcal{P}(\mathcal{X})$ of X with $|\mathcal{Q}| = m$. The hyperedge E_i^t is called *deme at generation t* of the elements of $Q_i \in \mathcal{Q}$. A dynamic population structure with $\mathcal{E}^t = \mathcal{E} = const$ for all $t \in \mathbb{N}$ is called *static population structure* oder just *population structure*.

Def. 6 The *diameter* of a dynamic population structure is $\min\{k \mid \sum_{i=0}^{k} \prod_{\nu=0}^{i} A_\nu(t) = K_\lambda\}$ with $A_\nu(t) = \overline{H_\nu^t}\ \overline{Q}$.

The pollination model from Eqn. 14 can now be extended by an isolation time η:

$$\begin{array}{lcl}\Pi_{\text{poll}} & = & (X, \mathcal{E}^t, \mathcal{Q}) \\ X & = & (0, \ldots, \lambda - 1) \\ Q_i & = & \{i\nu, \ldots, i\nu + \nu - 1\} \\ \mathcal{E}^t & = & (E_0^t, \ldots, E_{r-1}^t) \\ E_i^t & = & \begin{cases} Q_i \ \cup \ \bigcup_{(s,i)\in G} M_{s\to i}, \quad M_{s\to i} \subset Q_s & \text{for } t \equiv 0 \ mod \ \eta \\ Q_i & \text{else} \end{cases}\end{array} \tag{19}$$

7 Takeover Behavior of Population Structures

In the last years, a growing number of researchers in field of EAs try to find good measures for the selection pressure of different selection methods. Some of them are biologically inspired, e.g. [13, 14, 15], others base on pure probability theory, e.g. [16, 17, 18]. A decent analysis of properties of panmictic selection methods was given by Blickle and Thiele [19].

All these approaches relay on the assumption that individuals are indistinguishable and interchangeable. Since this does not hold true for structured populations, the theory developed for panmictic populations cannot be easily transfered.

7.1 Takeover Times and Probabilities

A common analytical approach to measure the selection pressure of an EA is the calculation of the takeover time, i.e. the number of generations it takes for the best individual of the initial population to fill the entire population under selection only. Unfortunately, for non-elitist selection operators, there is always a positive probability that the best individual gets lost before the population could be taken over, especially in the beginning.

An improvement of this approach is the notion of a takeover probability[18]. The idea is to define a Markov chain where each possible number of best individuals is an element of the state space.

Let be $p_{select} : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$, $(\lambda, k) \mapsto p_{select}(\lambda, k)$ a function calculating the success probability for a single trial if a population of λ individuals contains k best individuals, where success means drawing a best individual. Since selection of individuals is a Bernoulli experiment, the transition probability from a non-absorbing state i of the Markov chain into another state j is

$$p_{i,j} = \binom{\lambda}{j} (p_{select}(i, \lambda))^j (1 - p_{select}(i, \lambda))^{\lambda - j} \tag{20}$$

Because every trial has the same success probability, the success is binomial distributed. In the non-panmictic case, individuals are selected from different subsets of the population, thus trials have different success probabilities. Since this leads to a generalized binomial distribution, the state space of a Markov chain would be of size $\sum_{k=0}^{\lambda} \binom{\lambda}{k}$. Thus, a Markov chain analysis as in [18] is practically impossible.

7.2 Probabilistic Diameter

Although a simplification with respect to the extinction of the best individual, the takeover time as defined in [17] can be useful to compare selection methods among each other. The scenario for takeover time calculations is the following: The initial population contains a single best individual. Then, only selection is applied until the entire population consists of best individuals. The number of iterations needed is called the takeover time.

Section 6 contains the definition of the diameter of a population structure. Since the diameter is just the elitist takeover time of the population structure, it seems reasonable to calculate the takeover for non-elitist selection schemes in a similar manner. The idea is to propagate the best individual through the population as in the diameter calculation, based on the probability distribution of the selection operator. Chakraborty et.al. [18] calculated the success probabilities for the most common selection operators.

Starting with a single best individual in the initial population, the expectation value of the number of best individuals after one generation is

$$\mathsf{E}_{best}^{(1)} := \lambda \cdot p_{select}(\lambda, 1) \tag{21}$$

In general, for a given number $k^{(t)}$ of best individuals in generation t,

$$\mathsf{E}_{best}^{(t+1)} := \lambda \cdot p_{select}(\lambda, k^{(t)}) \tag{22}$$

is the expectation value for the number of best individuals in generation $t+1$. We obviously cannot calculate the expected number of best individuals in generation t by iterating eqn. 21 with

$$k^{(t+1)} := \mathsf{E}_{best}^{(t)} \tag{23}$$

This would mean to replace an iterated Bernoulli experiment by its expectation value. But then, this is how growth curve analysis and takeover time calculation was done in the past (e.g. [17], p. 71, eqn. 4).

To avoid the misleading notion of an expectation value, the following definition is given.

Def. 7 Let be $\Pi = (X, \mathcal{E}, \mathcal{Q}), |X| = \lambda$ a population structure, and $p_{select} : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}, (\lambda, k) \mapsto p_{select}(\lambda, k)$ a success function for a given selection operator. Let be for all $i \in X$:

$$s_i^1 = 1/\lambda \tag{24}$$

$$r_i^t = \sum_{j \in E_\nu} s_j^t, \quad i \in Q_\nu \tag{25}$$

$$s_i^{(t+1)} = p_{select}(r_i^t, |E_\nu|), \quad i \in Q_\nu \tag{26}$$

The *probabilistic diameter* of Π under p_{select} is

$$\Delta_\varepsilon(\Pi, p_{select}) := min\{t : \forall i \in X : s_i^{(t)} \geq 1 - \varepsilon\} \tag{27}$$

The value $1 - \varepsilon$ is the required takeover level, where ε is usually almost zero.

The recursive calculation of $\vec{r^t} = (r_0^t, \ldots, r_{\lambda-1}^t)$ can be done simultaneously:

$$\vec{r^t} = \vec{s}^{(t)\top} \overline{E}^{(t)}. \tag{28}$$

The value of $min\{s_i^{(t)} | i \in X\}$ can be interpreted as a growth coefficient of the best individual. Fig. 4 shows the growth rates for panmictic population models with population size 1024 for linear ranking selection as well as (μ, λ)-selection. As in [17], the growth curves resemble logistic functions, in contrast to the growth curves of non-panmictic population models.

The growth curves of the migration models in Fig. 5 and Fig. 6 show the different reaction of (μ, λ) selection and linear ranking to isolation times. The latter keeps the logistic character for much higher isolation times. The curve of the neighborhood model in Fig. 7 shows the expected linear growth rates.

8 Conclusions and Outlook

This paper presents a unified model of population structures in Evolutionary Algorithms. It has been shown that the model is a powerful framework for calculations on non-panmictic populations. As an example, we have presented a definition of the takeover time of arbitrary population structures which is consistent with the traditional panmictic takeover time definition.

This model presented here is intended as a base for further theoretical work in the field of non-panmictic population structures. While a Markov chain model of non-panmictic population structures is not generally useful, it should be possible to keep the state space small for a small number of demes by taking advantage of the duality of the given definition. This is subject to further research.

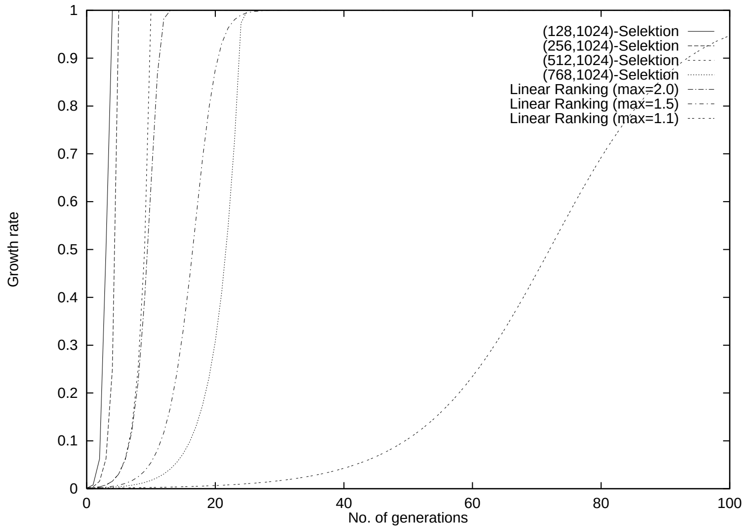

Figure 4: Growth curves of a panmictic (μ, λ) selection and linear ranking.

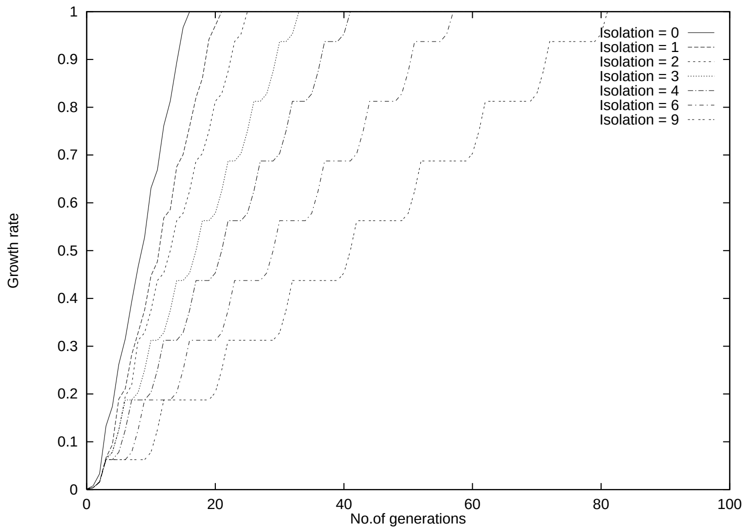

Figure 5: Growth curves of a migration model in a ring of 16 subpopulations and local (16,64) selection.

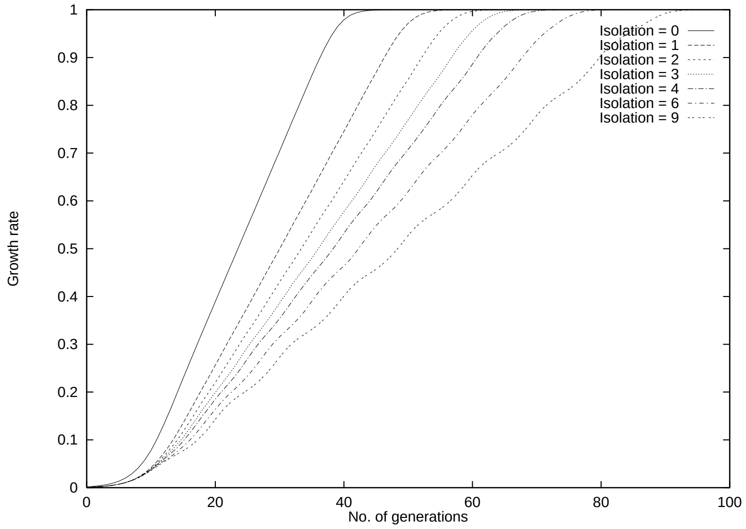

Figure 6: Growth curves of a migration model in a ring of 16 subpopulations and local linear ranking selection.

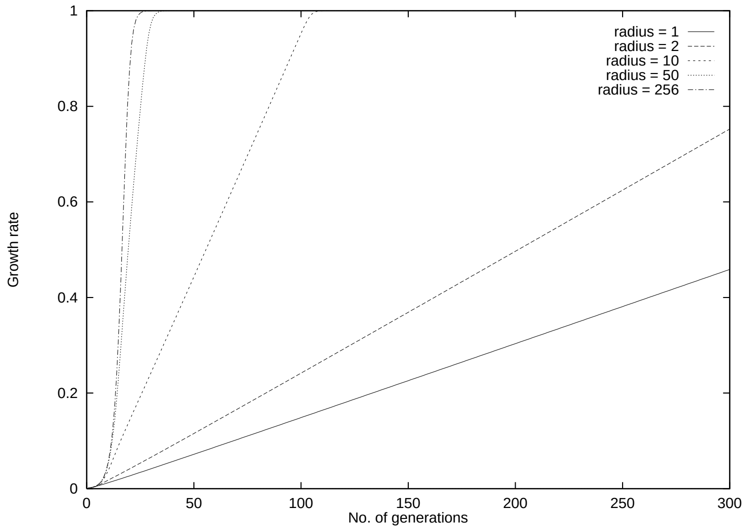

Figure 7: Growth curves of a neighborhood model in a ring of size 1024 and linear ranking selection (max = 1.5).

Acknowledgments

This work is a product of the Collaborative Research Center *Computational Intelligence* (SFB 531). Financial support by the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

References

[1] M. Gorges-Schleuter. *Genetic Algorithms and Population Structures — A Massively Parallel Algorithm.* Dissertation, Universität Dortmund, 1990.

[2] J. P. Cohoon, S. U. Hedge, W. N. Martin, and D. S. Richards. Punctuated Equilibria: A Parallel Genetic Algorithm. In Grefenstette [20], pages 148–154.

[3] Ch. C. Pettey, M. R. Leuze, and J. J. Grefenstette. A Parallel Genetic Algorithm. In Grefenstette [20], pages 155–161.

[4] R. Tanese. Parallel Genetic Algorithm for a Hypercube. In Grefenstette [20], pages 177–183.

[5] Günter Rudolph. Global optimization by means of distributed evolution strategies. In H.-P. Schwefel and R. Männer, editors, *Parallel Problem Solving from Nature – Proc. First Workshop PPSN*, pages 209–213, University of Dortmund, October 1–3, 1990. Springer, Berlin, 1991.

[6] Erick Cantu-Paz and David E. Goldberg. Predicting Speedups of Ideal Bounding Cases of Parallel Genetic Algorithms. In Thomas Bäck, editor, *Proceedings of the 7th International Conference on Genetic Algorithms*, pages 113–120, San Francisco, July19–23 1997. Morgan Kaufmann.

[7] Martina Gorges-Schleuter. ASPARAGOS An Asynchronous Parallel Genetic Optimization Strategy. In J. David Schaffer, editor, *Proceedings of the 3rd International Conference on Genetic Algorithms*, pages 422–427, George Mason University, June 1989. Morgan Kaufmann.

[8] Joachim Sprave. Linear neighborhood evolution strategy. In A. V. Sebald and L. J. Fogel, editors, *Proc. Third Annual Conf. Evolutionary Programming (EP'94)*, pages 42–51, San Diego CA, February 24–26, 1994. World Scientific, Singapore.

[9] Günter Rudolph and Joachim Sprave. Significance of locality and selection pressure in the grand deluge evolutionary algorithm. In H.-M. Voigt, W. Ebeling, I. Rechenberg, and H.-P. Schwefel, editors, *Parallel Problem Solving from Nature – PPSN IV, Int'l Conf. Evolutionary Computation*, pages 686–694, Berlin, 22.–26. September 1996. Springer, Berlin.

[10] J.A. Sarma. *An Analysis of Decentralized and Spatially Distributed Genetic Algorithms.* Dissertation, George Mason University, Fairfax, VA, 1998.

[11] Claude Berge. *Graphs and hypergraphs.* North-Holland mathematical library ; 6. North-Holland Publ. Comp., Amsterdam, 3. edition, 1973.

[12] Joachim Sprave. Zelluläre evolutionäre Algorithmen zur Parameteroptimierung. In R. Hofestädt, F. Krückeberg, and T. Lengauer, editors, *Informatik in den Biowissenschaften, 1. Fachtagung der GI-FG 4.0.2*, pages 111–120, Bonn, 15.–16. Februar 1993. Springer, Berlin.

[13] H. Mühlenbein and D. Schlierkamp-Voosen. Predictive Models for the Breeder Genetic Algorithm. *Evolutionary Computation*, 1(1):25–49, 1993.

[14] D. Thierens and D. E. Goldberg. Convergence Models of Genetic Algorithm Selection Schemes. In Y. Davidor, H.-P. Schwefel, and R. Männer, editors, *Parallel Problem Solving from Nature — PPSN III International Conference on Evolutionary Computation*, volume 866 of *Lecture Notes in Computer Science*, pages 119–129. Springer, Berlin, 1994.

[15] Th. Bäck. Generalized convergence models for tournament- and (μ,λ)-selection. In L. Eshelman, editor, *Proceedings of the 6th International Conference on Genetic Algorithms*, pages 2–8. Morgan Kaufmann Publishers, San Francisco, CA, 1995.

[16] A. E. Nix and M. D. Vose. Modeling genetic algorithms with Markov chains. *Annals of Mathematics and Artificial Intelligence*, 5:78–88, 1992.

[17] D. E. Goldberg and K. Deb. A Comparative Analysis of Selection Schemes Used in Genetic Algorithms. In G. J. E. Rawlins, editor, *Foundations of Genetic Algorithms*, pages 69–93. Morgan Kaufmann Publishers, San Mateo, CA, 1991.

[18] U. K. Chakraborty, Deb, K., and M. Chakraborty. Analysis of Selection Algorithms: A Markov Chain Approach. *Evolutionary Computation*, 4(2):133–167, 1996.

[19] T. Blickle and L. Thiele. A Comparison of Selection Schemes Used in Evolutionary Algorithms. *Evolutionary Computation*, 4(4):361–394, 1996.

[20] J. J. Grefenstette, editor. *Proceedings of the Second International Conference on Genetic Algorithms and Their Applications*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1987.