1 Introduction

One of the major challenges for researchers in the field of management science, information systems, business informatics, and computer science is to develop methods and tools that help organizations, such as companies or public institutions, to fulfill their tasks efficiently. However, during the last decade, the dynamics and size of tasks organizations are faced with has changed. Firstly, production and service processes must be reorganized in shorter time intervals and adapted dynamically to the varying demands of markets and customers. Although there is continuous change, organizations must ensure that the efficiency of their processes remains high. Therefore, optimization techniques are necessary that help organizations to reorganize themselves, to increase the performance of their processes, and to stay efficient. Secondly, with increasing organization size the complexity of problems in the context of production or service processes also increases. As a result, standard, traditional, optimization techniques are often not able to solve these problems of increased complexity with justifiable effort in an acceptable time period. Therefore, to overcome these problems, and to develop systems that solve these complex problems, researchers proposed using genetic and evolutionary algorithms (GEAs). Using these nature-inspired search methods it is possible to overcome some limitations of traditional optimization methods, and to increase the number of solvable problems. The application of GEAs to many optimization problems in organizations often results in good performance and high quality solutions.

For successful and efficient use of GEAs, it is not enough to simply apply standard GEAs. In addition, it is necessary to find a proper representation for the problem and to develop appropriate search operators that fit well to the properties of the representation. The representation must at least be able to encode all possible solutions of an optimization problem, and genetic operators such as crossover and mutation should be applicable to it.

Many optimization problems can be encoded by a variety of different representations. In addition to traditional binary and continuous string encodings, a large number of other, often problem-specific representations have been proposed over the last few years. Unfortunately, practitioners often report a significantly different performance of GEAs by simply changing the used representation. These observations were confirmed by empirical and theoretical investigations. The difficulty of a specific problem, and with it the performance of GEAs, can be changed dramatically by using different types of representations. Although it is well known that representations affect the performance of GEAs, no theoretical models exist which describe the effect of representations on the performance of GEAs. Therefore, the design of proper representations for a specific problem mainly depends on the intuition of the GEA designer and developing new representations is often a result of repeated trial and error. As no theory of representations exists, the current design of proper representations is not based on theory, but more a result of black art.

The lack of existing theory not only hinders a theory-guided design of new representations, but also results in problems when deciding which of the different representations should be used for a specific optimization problem. Currently, comparisons between representations are based mainly on limited empirical evidence, and random or problem-specific test function selection. However, empirical investigations only allow us to judge the performance of representations for the specific test problem, but do not help us in understanding the basic principles behind it. A representation can perform well for many different test functions, but fails for the one problem which one really wants to solve. If it is possible to develop theoretical models which describe the influence of representations on measurements of GEA performance – like time to convergence and solution quality – then representations can be used efficiently and in a theory-guided manner. Choosing and designing proper representations will not remain the black art of GEA research but become a well predictable engineering task.

1.1 Purpose

The purpose of this work is to bring some order into the unsettled situation which exists and to investigate how representations influence the performance of genetic and evolutionary algorithms. This work develops elements of representation theory and applies them to designing, selecting, using, choosing among, and comparing representations. It is not the purpose of this work to substitute the current black art of choosing representations by developing barely applicable, abstract, theoretical models, but to formulate an applicable representation theory that can help researchers and practitioners to find or design the proper representation for their problem. By providing an applicable theory of representations this work should bring us to a point where the influence of representations on the performance of GEAs can be judged easily and quickly in a theory-guided manner.

The first step in the development of an applicable theory is to identify which properties of representations influence the performance of GEAs and how. Therefore, this work models for different properties of representations how solution quality and time to convergence is changed. Using this theory, it is possible to formulate a framework for efficient design of representations. The framework describes how the performance of GEAs, measured by run duration and solution quality, is affected by the properties of a representation. By using this framework, the influence of different representations on the performance of GEAs can be explained. Furthermore, it allows us to compare representations in a theory-based manner, to predict the performance of GEAs using different representations, and to analyze and design representations guided by theory. One does not have to rely on empirical studies to judge the performance of a representation for a specific problem, but can use existing theory for predicting GEA performance. By using this theory, the situation exists where empirical results are only needed to validate theoretical predictions.

However, developing a general theory of how representations affect GEA performance is a demanding and difficult task. To simplify the problem, it must be decomposed, and the different properties of encodings must be investigated separately. Three different properties of representations are considered in this work: Redundancy, scaling, and locality, respectively distance distortion. For these three properties of representations models are developed that describe their influence on the performance of GEAs. Additionally, population sizing and time to convergence models are presented for redundant and non-uniformly scaled encodings. Furthermore, it is shown that low-locality representations can change the difficulty of the problem. For low-locality encodings, it can not exactly be predicted how GEA performance is changed, without having complete knowledge regarding the structure of the problem. Although the investigation is limited only to three important properties of representations, the understanding of the influence of these three properties of encodings on the performance of GEAs brings us a large step forward towards a general theory of representations.

To illustrate the significance and importance of the presented representation framework on the performance of GEAs, the framework is used for analyzing the performance of binary representations of integers and tree representations. The investigations show that the current framework considering only three representation properties gives us a good understanding of the influence of representations on GEA performance as it allows us to predict the performance of GEAs using different types of representations. The results confirm that choosing a proper representation has a large impact on the performance of GEAs, and therefore, a better theoretical understanding of representations is necessary for an efficient use of genetic search.

Finally, it is illustrated how the presented theory of representations can help us in designing new representations more reasonably. It is shown by example for tree representations, that the presented framework allows theory-guided design. Not black art, but a deeper understanding of representations allows us to develop representations which result in a high performance of genetic and evolutionary algorithms.

1.2 Organization

The organization of this work follows its purpose. It is divided into two large parts: After the first two introductory chapters, the first part (Chaps. 3 and 4) provides the theory regarding representations. The second part (Chap. 5, 6, 7, and 8) applies the theory to the analysis and design of representations. Chapter 3 presents theory on how different properties of representations affect GEA performance. Consequently, Chap. 4 uses the theory for formulating the time-quality framework. Then, in Chap. 5, the presented theory of representations is used for analyzing the performance differences between binary representations of integers. Finally, the framework is used in Chap. 6, Chap. 7, and Chap. 8 for the analysis and design of tree representations and search operators. The following paragraphs give a more detailed overview about the contents of each chapter.

Chapter 1 is the current chapter. It sets the stage for the work and describes the benefits that can be gained from a deeper understanding of representations for GEAs.

Chapter 2 provides the background necessary for understanding the main issues of this work about representations for GEAs. Section 2.1 introduces representations which can be described as a mapping that assigns one or more genotypes to every phenotype. The genetic operators selection, crossover, and mutation are applied on the level of alleles to the genotypes, whereas the fitness of individuals is calculated from the corresponding phenotypes. Section 2.2 illustrates that selectorecombinative GEAs, where only crossover and selection operators are used, are based on the notion of schemata and building blocks. Using schemata and building blocks is an approach to explain why and how GEAs work. This is followed in Sect. 2.3 by a brief review of reasons and measurements for problem difficulty. Measurements of problem difficulty are necessary to be able to compare the influence of different types of representations on the performance of GEAs. The chapter ends with some earlier, mostly qualitative recommendations for the design of efficient representations.

Chapter 3 presents three aspects of a theory of representations for GEAs. It investigates how redundant encodings, encodings with exponentially scaled alleles, and representations that modify the distances between the corresponding genotypes and phenotypes, influence GEA performance. Population sizing models and time to convergence models are presented for redundant and exponentially scaled representations. Section 3.1 illustrates that redundant encodings influence the supply of building blocks in the initial population of GEAs. Based on this observation the population sizing model from Harik et al. (1997) and the time to convergence model from Thierens and Goldberg (1993) can be extended from non-redundant to redundant representations. Because redundancy mainly affects the number of copies in the initial population that are given to the optimal solution, redundant representations increase solution quality and reduce time to convergence if individuals that are similar to the optimal solution are overrepresented. Section 3.2 focuses on exponentially scaled representations. The investigation into the effects of exponentially scaled encodings shows that, in contrast to uniformly scaled representations, the dynamics of genetic search are changed. By combining the results from Harik et al. (1997) and Thierens (1995) a population sizing model for exponentially scaled building blocks with and without considering genetic drift can be presented. Furthermore, the time to convergence when using exponentially scaled representations is calculated. The results show that when using non-uniformly scaled representations, the time to convergence increases. Finally, Sect. 3.3 investigates the influence of representations that modify the distances between corresponding genotypes and phenotypes on the performance of GEAs. When assigning the genotypes to the phenotypes, representations can change the distances between the individuals. This effect is denoted as locality or distance distortion. Investigating its influence shows that the size and length of the building blocks, and therefore the complexity of the problem are changed if the distances between the individuals are not preserved. Therefore, to ensure that an easy problem remains easy, high-locality representations which preserve the distances between the individuals are necessary.

Chapter 4 presents the framework for theory-guided analysis and design of representations. The chapter combines the three elements of representation theory from Chap. 3 – redundancy, scaling, and locality – to a time-quality framework. It formally describes how the time to convergence and the solution quality of GEAs depend on these three aspects of representations. The chapter ends with implications for the design of representations which can be derived from the framework. In particular, the framework tells us that uniformly scaled representations are robust, that exponentially scaled representations are fast but inaccurate, and that low-locality representations change the difficulty of the underlying optimization problem.

Chapter 5 uses the framework for a theory-guided analysis of binary representations of integers. Because the potential number of schemata is higher when using binary instead of integer representations, users often favor the use of binary instead of integer representations, when applying GEAs to integer problems. By using the framework it can be shown that the redundant unary encoding results in low GEA performance if the optimal solution is underrepresented. Both, Gray and binary encoding are low-locality representations as they change the distances between the individuals. Therefore, both representations change the complexity of optimization problems. It can be seen that the easy integer one-max problem is easier to solve when using the binary representation, and the difficult integer deceptive trap is easier to solve when using the Gray encoding.

Chapter 6 uses the framework for the analysis and design of tree representations. For tree representations, standard crossover and mutation operators are applied to tree-specific genotypes. However, finding or defining tree-specific genotypes and genotype-phenotype mappings is a difficult task because there are no intuitive genotypes for trees. Therefore, researchers have proposed a variety of different, more or less tricky representations which can be used in combination with standard crossover and mutation operators. A closer look at the Prüfer number representation in Sect. 6.2 reveals that the encoding in general is a low-locality representation and modifies the distances between corresponding genotypes and phenotypes. As a result, problem complexity is modified, and many easy problems become too difficult to be properly solved using GEAs. Section 6.3 presents an investigation into the characteristic vector representation. Because invalid solutions are possible when using characteristic vectors, an additional repair mechanism is necessary which makes the representation redundant. Characteristic vectors are uniformly redundant and GEA performance is independent of the structure of the optimal solution. However, the repair mechanism results in non-synonymous redundancy. Therefore, GEA performance is reduced and the time to convergence increases. With increasing problem size, the repair process generates more and more links randomly and offspring trees have not much in common with their parents. Therefore, for larger problems guided search is no longer possible and GEAs behave like random search. In Sect. 6.4, the investigation into the redundant link and node biased representation reveals that the representation overrepresents trees that are either star-like or minimum spanning tree-like. Therefore, GEAs using this type of representation perform very well if the optimal solution is similar to stars or to the minimum spanning tree, whereas they fail when searching for optimal solutions that do not have much in common with stars or the minimum spanning tree. Finally, Sect. 6.5 presents network random keys (NetKeys) as an example for the theory-guided design of a tree representation. To construct a robust and predictable tree representation, it should be non- or uniformly redundant, uniformly scaled, and have high-locality. When combining the concepts of the characteristic vector representation with weighted representations like the link and node biased representation, the NetKey representation can be created. In analogy to random keys, the links of a tree are represented as floating numbers, and a construction algorithm constructs the corresponding tree from the keys. The NetKey representation allows us to distinguish between important and unimportant links, is uniformly redundant, uniformly scaled, and has high locality.

Chapter 7 uses the insights into representation theory for the analysis and design of search operators for trees. In contrast to Chap. 6 where standard search operators are applied to tree-specific genotypes, now tree-specific search operators are directly applied to tree structures. Such types of representations are also known as direct representations as there is no additional genotype-phenotype mapping. Section 7.1 presents a direct representation for trees (NetDir) as an example for the design of direct tree representations. Search operators are directly applied to trees and problem-specific crossover and mutation operators are developed. The search operators for the NetDir representation are developed based on the notion of schemata. Section 7.2 analyzes the edge-set encoding which encodes trees directly by listing their edges. Search operators for edge-sets may be heuristic, considering the weights of edges they include in offspring, or naive, including edges without regard to their weights. Analyzing the properties of the heuristic variants of the search operators shows that solutions similar to the minimum spanning tree are favored. In contrast, the naive variants are unbiased which means that genetic search is independent of the structure of the optimal solution. Although no explicit genotype-phenotype mapping exists for edge-sets and the framework for the design of representations cannot be directly applied, the framework is useful for structuring the analysis of edge-sets. Similarly to non-uniformly redundant representations, edge-sets overrepresent some specific types of tree and GEA performance increases if optimal solutions are similar to the MST. Analyzing and developing direct representations nicely illustrates the trade-off between designing either problem-specific representations or problem-specific operators. For efficient GEAs, it is necessary either to design problem-specific representations and to use standard operators like one-point or uniform crossover, or to develop problem-specific operators and to use direct representations.

Chapter 8 verifies theoretical predictions concerning GEA performance by empirical verification. It compares the performance of GEAs using different types of representations for the one-max tree problem, the deceptive tree problem, and various instances of the optimal communication spanning tree problem. The instances of the optimal communication spanning trees are presented in the literature (Palmer 1994; Berry et al. 1997; Raidl 2001; Rothlauf et al. 2002). The results show that with the help of the framework the performance of GEAs using different types of representations can be well predicted.

Chapter 9 summarizes the major contributions of this work, describes how the knowledge about representations has changed, and gives some suggestions for future research.