

Improving Robustness in Social Fabric-based Cultural Algorithms

By

Bahram Zaeri

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science

in the School of Computer Science

© Bahram Zaeri, 2017

University of Windsor

All rights reserved. This dissertation may not be reproduced in whole or in part, by
photocopying or other means, without the permission of the author.

Improving Robustness in Social Fabric-based Cultural Algorithms

By

Bahram Zaeri

Supervisory Committee

Dr. Ziad Kobti, Supervisor
(School of Computer Science)

Dr. Mehdi Kargar, Internal Reader
(School of Computer Science)

Dr. Kemal Tepe, External Reader
(Department of Electrical and Computer Engineering)

ABSTRACT

In this thesis, we propose two new approaches which aim at improving robustness in social fabric-based cultural algorithms. Robustness is one of the most significant issues when designing evolutionary algorithms. These algorithms should be capable of adapting themselves to various search landscapes.

In the first proposed approach, we utilize the dynamics of social interactions in solving complex and multi-modal problems. In the literature of Cultural Algorithms, Social fabric has been suggested as a new method to use social phenomena to improve the search process of CAs. In this research, we introduce the Irregular Neighborhood Restructuring as a new adaptive method to allow individuals to rearrange their neighborhoods to avoid local optima or stagnation during the search process.

In the second approach, we apply the concept of Confidence Interval from Inferential Statistics to improve the performance of knowledge sources in the Belief Space. This approach aims at improving the robustness and accuracy of the normative knowledge source. It is supposed to be more stable against sudden changes in the values of incoming solutions.

The IEEE-CEC2015 benchmark optimization functions are used to evaluate our proposed methods against standard versions of CA and Social Fabric. IEEE-CEC2015 is a set of 15 multi-modal and hybrid functions which are used as a standard benchmark to evaluate optimization algorithms. We observed that both of the proposed approaches produce promising results on the majority of benchmark functions. Finally, we state that our proposed strategies enhance the robustness of the social fabric-based CAs against challenges such as multi-modality, copious local optima, and diverse landscapes.

Contents

Supervisory Committee	ii
Abstract	iii
Table of Contents	iv
List of Tables	vii
List of Figures	viii
Acknowledgements	ix
Dedication	x
1 Introduction	1
1.1 Evolutionary Algorithms	1
1.2 Research Motivation	2
1.3 Thesis Contribution	3
1.4 Thesis Outline	3
2 Related Works	4
2.1 Cultural Algorithms	4
2.1.1 Heterogeneous Multi-Population Cultural Algorithm	4
2.1.2 The Social Fabric Approach as an Approach to Knowledge In- tegration in Cultural Algorithms	5

2.1.3	Robust Evolution Optimization at the Edge of Chaos: Commercialization of Culture Algorithms	7
2.1.4	Socio-Cultural Evolution via Neighborhood-Restructuring in Intricate Multi-Layered Networks	8
2.1.5	Leveraged Neighborhood Restructuring in Cultural Algorithms for Solving Real-World Numerical Optimization Problems . .	10
2.2	Particle Swar Optimization	12
2.2.1	Tribe-PSO: A novel global optimization algorithm and its application in molecular docking	12
2.2.2	Heterogeneous Particle Swarm Optimizers	13
3	Evolutionary Algorithms	16
3.1	Evolutionary Algorithms	16
3.2	Cultural Algorithms	18
3.2.1	Belief Space	18
3.2.2	Population Space	19
3.2.3	Communication Protocol	19
3.3	Social Fabric-based Cultural Algorithms	20
3.3.1	Evolution Phases	20
3.3.2	Strategic Neighborhood Restructuring	21
3.3.3	Social fabric based influence function	21
3.4	Particle Swarm Optimization	22
3.4.1	Tribe-PSO	23
4	Proposed Approaches: Improving Robustness in Social Fabric	24
4.1	Self-organization	25
4.2	Irregular Neighborhood Restructuring	26
4.3	Confidence based Normative Knowledge Source	27

5	Experimental Setup	28
5.1	Description	28
5.2	Benchmark Functions	29
5.2.1	Unimodal Functions	29
5.2.2	Simple Multimodal Functions	30
5.2.3	Hybrid Functions	32
5.2.4	Composite Functions	33
5.3	Parameters Settings	35
A	Additional Information	37
	Bibliography	38

List of Tables

List of Figures

ACKNOWLEDGEMENTS

I would like to thank:

my cat, Star Trek, and the weather, for supporting me in the low moments.

Supervisor Main, for mentoring, support, encouragement, and patience.

Grant Organization Name, for funding me with a Scholarship.

I believe I know the only cure, which is to make one's centre of life inside of one's self, not selfishly or excludingly, but with a kind of unassailable serenity-to decorate one's inner house so richly that one is content there, glad to welcome any one who wants to come and stay, but happy all the same in the hours when one is inevitably alone.

Edith Wharton

DEDICATION

Just hoping this is useful!

Chapter 1

Introduction

1.1 Evolutionary Algorithms

Evolutionary algorithms take their roots from the evolution theory of Darwin. These algorithms try to mimic the collective behavior of natural systems. Natural processes such as natural selection, survival of the fittest, and reproduction have been subject to inspiration as the fundamental components of evolutionary problem-solving methods. The basic part of all these algorithms is that they start with a randomly generated set of solutions. Then, they try to evolve the solutions through applying a set evolutionary operators such as mutation and crossover.

Evolutionary algorithms are not limited to biological processes. There is another category of evolutionary algorithms known as Swarm Intelligence (SI). These algorithms take inspiration from social behaviors of living colonies such as ants, flocks, schools, and hives. Within these swarms, individuals have relatively simple structures, but their collective behavior usually looks very complex. The complex behavior of a swarm emerges as a result of the interactions between the individuals over time. This complex behavior can not be easily predicted by observing the simple behaviors of the agents separately.

There are some well-known examples categorized as evolutionary algorithms:

1. Genetic Algorithms
2. Cultural Algorithms
3. Particle Swarm Optimization
4. Ant Colony Optimization
5. Honey Bee Colony

Both categories of evolutionary algorithms share the same idea of evolving some initially generated solutions. However, the difference is in the way that they manipulate and evolve individuals through applying evolutionary operators.

1.2 Research Motivation

The aim of this research work is to improve robustness in evolutionary algorithms. In this field, robustness means that an algorithm can be used to solve many kinds of problems, with a minimum number of adjustments to address particular problems with special qualities. Also, it can mean the capability of algorithms to deal acceptably with noisy or missing data.

As stated by No Free Lunch(NFL) theorem, there is no algorithm better than others over all cost functions. It means, there is no guarantee for an algorithm to work well for all functions if it shows promising results for a particular category of them. Therefore, robustness has been one of the most desired features which motivates researchers to invent new methods which are less dependent on the kind of a problem than others. In our thesis, we mostly emphasize on exploring different approaches to improve robustness in cultural algorithms regarding both belief and population spaces.

1.3 Thesis Contribution

In the thesis, I am going to improve the robustness of Cultural Algorithms in both components of population and belief spaces. In the population space, a new neighborhood restructuring strategy is proposed which works based on a dynamic and irregular topology. In our approach, neighborhood restructuring occurs at a microscopic level, and every individual decides to change its neighborhood size. In the belief space, a new normative knowledge source is proposed which works based on the Confidence Interval concept inspired from Inferential Statistics. Also, we recognized that both Social fabric and PSO algorithm utilize the same social structure to facilitate social interactions between the individuals. So, we decided to use PSO in the population component of our proposed approach. We hypothesize that both of our approaches help the algorithm to resist against perturbations and not get affected by the fluctuations in the search space. The first approach lets the individuals adjust their relationships autonomously (self-organization). The second method makes the normative ranges in the belief space robust against fluctuations in the input data.

Various benchmark functions have been used to evaluate the efficiency of evolutionary algorithms. As a reliable benchmark, we chose IEEE-CEC2015 function set which covers most of the desired properties for an optimization testbed such as multi-modality, copious local optima, and non-separability. In our thesis, they are categorized in four categories: Unimodal, Simple multimodal, Hybrid, and Composite functions.

1.4 Thesis Outline

abcd

Chapter 2

Related Works

Many variants of CAs have been proposed in a vast range of different applications such as single and multi-objective optimization, dynamic problems, social interactions simulation. Here, we are interested in studying socially motivated and multi-population variants as some modern approaches in solving optimization problems.

2.1 Cultural Algorithms

2.1.1 Heterogeneous Multi-Population Cultural Algorithm

[1] proposed a new architecture for cultural algorithms. In this approach, the whole population is divided into a set of independent sub-populations which work in parallel without direct communication. They referred to the works of [2, 3, 4]. As their motivation they stated that most of proposed variants of evolutionary algorithms suffers from immature convergence. This occurs because these algorithms can not hold the diversity at a reasonable level. Based on existing research works, they hypothesized that multi-population strategies would be a better choice as they have the potential to perform an efficient search on complex landscapes. In their approach, the optimization parameters are divided among some heterogeneous sub-populations.

the sub-populations are called heterogenous because each sub-population is responsible for optimizing a different subset of parameters. Each sub-population represents a partial solution instead of a complete solution. To evaluate a partial solution, it gets completed by its complement parameter values from the belief space. The complete solution is evaluated based on a numerical optimization function. Also, to make the convergence process faster a simple local search strategy is incorporated into the proposed algorithm. The general architecture of their algorithm is presented in figure ???

In the experiments, they considered the whole population size to be 1000 individuals. It is divided into 30 sub-populations. So, the size of each sub-population is 33. The algorithm runs for the maximum of 10000 generations and the local search strategy runs only for 10 iterations. They evaluated HMP-CA on a set of 8 complex optimization functions. It is able to find the minimum value for 7 functions out of 8. However, when the local search strategy is applied to the experiments, the proposed method outperforms all of the functions and it finds the optimum value very quickly. Ultimately, they claimed that their proposed approach is efficient in both time and space complexity.

2.1.2 The Social Fabric Approach as an Approach to Knowledge Integration in Cultural Algorithms

[] begins with a brief introduction to socially motivated methods to problem solving. It compares qualities of Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), and Cultural Algorithms (CA) regarding the scale in which the interactions between agents occur. Figure (()) compares PSO, ACO, and CA in terms of the time and space continuum over which the social interactions occur. Individuals in ACO and PSO tend to interact in a relatively limited temporal and spatial scales. It is obvious because the agents in both ant and particle swarm algorithms exchange

information with only other agents in their local neighborhood. On the other hand, cultural algorithms let the individuals interact together using various types of symbolic information emerged from complex cultural systems. In cultural algorithms the interactions among individuals occur indirectly through a shared belief space. So, cultural algorithms allow individuals interact in a global scale.

Then, they asked the essential question of what social structures might emerge alongside the search process?. To answer such questions, they introduced a new influence function which utilizes the social fabric phenomena. The old influence function assumes no interactions between agents and works based on the simple roulette wheel method. On the other hand, in the new influence function, the individuals are connected through a social network (fabric). Multiple layers of such networks could be employed in a population. The interplay of these network connection forms a social fabric. At each iteration, an individual could specify its controller knowledge source. In this approach, the controller knowledge source is chosen based on the majority of knowledge sources in the neighborhood of an individual. The neighborhood size of an individual is specified by the topology of the fabric. Inspired by Particle Swarm Optimization literature, different topologies could be taken into consideration to model the relationships among individuals. In their work, they only considered Ring and Square topologies. They stated that, the topology of the social fabric determines the extent to which the influence of knowledge sources could be spread thorough the network.

To evaluate the social fabric approach, they implemented it in Repast framework. Repast is a simulation tool for multi-agent systems. They created a cultural algorithms toolkit (CAT) to view the capabilities of cultural algorithms in solving various problems. They chose Cone World problem to evaluate and compare their approach with the standard cultural algorithm. The reason that they chose this problem is that by changing its parameters during the evolution process, it can show a dynamic be-

havior. So, Cone World problem provides an efficient way to test flexibility of search algorithms. They set the parameters of CAT as: 100 individuals, 100 cones and 1000 generations. They used ring and square topologies to form the social fabric. They stated that square topology works better than ring as it finds the solution after 250 iterations. While, the ring topology finds the best solution 450 iterations.

2.1.3 Robust Evolution Optimization at the Edge of Chaos: Commercialization of Culture Algorithms

The authors of [1] aimed at commercializing Cultural Algorithm Toolkit (CAT) through developing a robust variant of it. By robustness, they mean to develop a cultural algorithm which is capable of being applied across a vast range of complex problems. At first, they referred to [Peng model] as a standard model cultural algorithms which assumes no connection between individuals. Then, they referred to [Ali], which introduced the concept of social fabric to allow individuals interact together. The authors extended the work of Ali by allowing the social networks having a memory. In addition, they utilized a variety of different networks in order to determine the relationship between network and problem complexity.

They brought up the hypothesis that there might be multiple independent networks for different purposes such as kinship and economics. In such networks, there are always some individuals which are member of multiple networks and so, they can play the role of mediator between different networks. To preserve the diversity, the authors utilized a variety of different topologies such as Lbest, Square, Hexagon, Octagon, Hexadecagon, and Gbest.

They stated that in the previous work of [Ali], he employed an un-weighted majority win strategy to determine the controller knowledge source of an individual. It just relied on the count of each knowledge source in the neighborhood of each individual. The authors replaced it with a new strategy which uses the average fitness of each

KSs instead of each KSs count. So, the performance of individuals in a neighborhood influences which knowledge source to be chosen for the next iteration.

To evaluate the robustness of the algorithm, the authors used Cone World Generator [1] as a dynamic problem. Referred to the work by [Langton], they defined three classes of entropy in the cones world problem:

1. Fixed: There is a low entropy and the parameters of the environment do not tend to change at all.
2. Periodic: The environment parameters change in a regular period of time. So, the algorithm should be capable adapting itself regularly.
3. Chaotic: In this case, the parameters change without any order and no prediction could be made about them. In fact, it is more similar to real-world cases.

For the fixed category, the square topology successfully solved 84% of the problems and as the best topology. For the periodic case, the octagon topology showed a better performance than other topologies. And, for the chaotic category, Gbest showed a better result mostly in terms of average time to find a solution and the standard deviation. In summary, as the complexity of problems grows, there is the need for more connections between individuals in a social fabric to keep the robustness of the algorithm.

2.1.4 Socio-Cultural Evolution via Neighborhood-Restructuring in Intricate Multi-Layered Networks

The authors of [1] aimed at exploring the utilization of neighborhoods in the population level of cultural algorithms and see how it influences the knowledge swarming in the belief space. Their approach uses a dynamic neighborhood restructuring to preserve

diversity efficiently. They referred to the reasearch works of [1], [2], [3]. Those papers proposed some adaptive methods which tried to make a trade-off between exploration and eploitation properties of evolutionary algoithms. Among the referred papers, the work by [ALi] used the social fabric phenomena to from multi-layered hierarchical network structures in both homogenous and heterogeneous networks.

The authors defined the social fabric phenomena as follows:

”The Social Fabric is a living informational skin created out of the engineered emergence of agents illustrating the tension between the individual and the community in a context of interaction between them”

The informational skin is created by the connectivity of individuals together. The social fabric is used to combine the behaviors at both indivials and society levels. Then, they proposed a strategy to determine the contoller knowledge source of each individual in each generation. In the strategy, the agents send the name of their contoller knowledge source to their neighbord through the social fabric. Each individual picks the mostly used knowledge source in its neighborhood as its controller for the next iteration. In case of a tie between two or more knowledge sources, three tie-breaking strategies are intorduced:

1. Most Frequently Used (MFU)
2. Random
3. Least Frequently Used (LFU)

They stated that their proposed approach is inspired from a previous work of Ali[22]. The new idea was called Cultual Algorithm with Restructuring Layered Social Fabric (CARLSOF). In this approach, the whole population is divided into two layers of rudimentary and advanced members. There are some independent tribes in the population space which their best performing individuals form the advanced layer and

the others form the rudimentary level. In the previous works, the topology of social fabrics were supposed to be fixed and homogenous. They utilized a variety of regular topologies to form the social fabric. The authors proposed a new strategy to enable the fabric to be restructured in the case of stagnation. Each tribe may decide to change its topology to a denser (with more connections) or sparser (with less connections). They referred to two different restructuring strategies as similar approaches. The first one was Layered Delaunay Triangulation which is based on voronoi diagram. The second approach was Random Rewiring Procedure which starts with a regular topology and then rewires each edge randomly with the probability p .

To evaluate the performance of CARLSOF, they used function set of IEEE-CEC2011 evolutionary competition. The algorithm was implemented in JAVA. The authors reported that CARLSOF was successful to enhance all of the functions in the testbed in terms of average, best, and worst obtained values. Also, they claimed that CARLSOF outperforms the best previously obtained results for European Space Agency (P12) and Casini (P13) problems results. They reported 2.983 km/sec compared to previous best 7.095. For p13, they reported 8.383091 km/sec compared to previous best 8.3832.

2.1.5 Leveraged Neighborhood Restructuring in Cultural Algorithms for Solving Real-World Numerical Optimization Problems

In the paper, the authors made an overall review of related works on Cultural Algorithms and Social Fabric [1], [2], [3]. They introduced the concept of social fabric as an infrastructure which facilitates the propagation of the influence of the knowledge sources through the population space. The whole population is divided into multiple small sized tribes. Unlike other PSO-based multi-swarm variants, the tribal sub-

swarms change their topology during the evolution process to keep diversity against stagnations.

The tribes are organized in a two-layer structure of advanced and rudimentary classes. The advanced class is composed of best performing individuals of each tribe. From each tribe, two exemplars are chosen. The first one is the best of explorer knowledge sources (Normative and Topographic) and the second one is the best of exploiter knowledge sources (Situational). The authors utilized only three kinds of knowledge sources: Normative, Topographic, and Situational. They claimed that, other types of knowledge sources are suitable for dynamic problems. However, they used a set of static problems to evaluate their approach. The authors divided the evolution process of CAs into three stages: the seclusion stage, the rapport stage, and the cohesive stage. In the seclusion stage, the tribes evolve independently and without any exchange of information. Then, in the rapport stage, the tribes start to exchange knowledge via the advanced class. Also, the neighborhood restructuring strategy is launched in each tribe. So, the tribes have the capability to restructure their topology to keep diversity and avoid local optima. The first two stages ensure that diversity has been kept in the initial iterations of the search process. Finally, in the third stage, all the tribes are merged together and continue the search process as a single CA model.

Based on the categorization proposed in [1], the authors claimed that their approach is heterogeneous regarding update-rule heterogeneity and neighborhood heterogeneity. Since the topology of each tribe is changing during the evolution process, the tribes might use different topologies at the same time. So, it could be considered as heterogeneous neighborhood restructuring. Also, their approach utilizes update-rule heterogeneity because, there are individuals with different roles (explorer and exploiter).

In the experimental results section, the authors used IEEE-CEC2011 competition on testing evolutionary algorithms on real-world numerical optimization problems. They

presented an extended comparison of their approach with other proposed methods. At first, the authors tested the approach (T-SCANeR) with different values for these parameters: tie-breaking rule, M_{thresh} , and number of elites. There is no best value for tie-breaking rule as it is problem-dependent. For the M_{thresh} parameter, the results show that the best value is 50. The best value for n_{tElite} is 2. The second experiment was on varying the window size within which the social fabricis applied and aggregated (W_{size}). The best reported value for W_{size} is 20.

In the two first experiments the best values for the parameters were determined. Based on the obtained values, the authors compared T-SCANeR with five other evolutionary algorithms on the IEEE-CEC2011 which is a set of 18 functions. T-SCANeR managed to outperform for most of the functions except for T_3 , T_4 , T_6 , and $T_{11.5}$. Problems T_{12} (full messenger problem) and T_{13} (cassini problem) are associated with the European Space Agency (ESA). The authors claimed that T-SCANeR outperformed other methods by 2.193724 km/s for the messenger problem and 8.383090 for the cassini problem.

2.2 Particle Swar Optimization

2.2.1 Tribe-PSO: A novel global optimization algorithm and its application in molecular docking

The authors proposed a new variant of Particle Swarm Optimization (PSO) called Tribe-PSO for the primary purpose of molecular docking in chemometrics. As some previous research work, they referred to [1], [2], [3], [4]. Their approach is inspired from Hierarchical Fair Competition concept [5]. They divided the whole population into some tribes and the evolution process into three phases. The tribes are organized into two layers of basic and upper individuals.

The individuals in each tribe are completely isolated from other tribes. So, they

form the basic layer. On the other hand, the best members of each tribe can see the other tribes and exchange information with their best members as well. The problem-solving process is divided into three phases: isolated, communing, and united. In the first phase, the tribes are completely isolated and there is no exchange of information. In the second phase, the two-layered model is formed and the tribes begin to exchange information through their best performing individuals. In the third phase, all the tribes are merged into a single population. Then, it operates as basic PSO model until meeting some stopping criteria. The authors claimed that their approach helps the individuals preserving diversity and avoiding local optima against multi-modal solution spaces.

The authors used two testbeds to evaluate and compare the performance of Tribe-PSO with standard PSO. The first was De Jong's function set and the second was a test set of 100 receptor-ligand X-ray structures selected for docking benchmark. In the De Jong's testbed, the basic PSO showed a better performance in the isolated phase. However, in the communing phase Tribe-PSO converged to a better value than basic PSO. In the unity phase, the basic PSO seemed to get stuck in a local optimum. However, Tribe-PSO continued to accelerate convergence and got better results than basic PSO. In the docking benchmark, Tribe-PSO was compared to the AutoDock library. Four parameters are calculated after 10 independent runs: Best, Run1, Average, and Standard Deviation of the results. The relative difference of the four benchmark factors between Tribe-PSO and AutoDock were calculated. The results for the four factors showed that Tribe-PSO leads to a better performance than AutoDock.

2.2.2 Heterogeneous Particle Swarm Optimizers

The authors presented a survey on heterogeneous variants of Particle Swarm Optimization (PSO). They claim that the homogenous models have been attractive

because of their simplicity in conceptual and application levels. However, heterogeneous models are ubiquitous in nature. Here, heterogeneity means the individuals may differ from each other regarding their parameters and search behavior.

At first, the authors presented a brief description of three well-known variants of PSO: Accelerated PSO[ref], Fully-informed PSO[ref], and Bare Bones PSO[ref]. Then, they authors categorized the heterogeneous variants of PSO into four categories:

1. Neighborhood
2. Model of Influence
3. Update Rule
4. Parameters

Neighborhood heterogeneity refers to the cases that the topology of the swarm is not regular. This type of heterogeneity occurs when the neighborhood size of each individual is different. They claimed that the neighborhood heterogeneity allows some population to be more influential than others. Individuals with higher number of connections have the potential to attract more individuals through the search process. [Kennedy ref]

Model of Influence heterogeneity refers to the situations that the individuals employ different strategies to specify their informers. The word Informer refers to an individual which is going to influence another individual. []

Update-rule heterogeneity means the individuals utilize different strategies to update their position in the search space. This kind of heterogeneity makes it possible for the individuals to explore the search space in several ways. Also, the particles can play different but complementary roles. For example, some of them may tend to explore unseen parts of the solution space and some others only follow those scout individuals. The second type are the exploiters.

Parameter heterogeneity occurs when some individuals in a group which follow the same update rule use different update rule's parameter settings. Having different search parameters, even in a group of similar particles, leads to various search behaviors which improves the level diversity. The authors referred to [1, 2, 3] which utilize different initial values for the parameters such as acceleration coefficient, maximum velocities and inertia weight.

To compare the mentioned categories, the authors created two test cases. The first one compared two PSO variants with different update rules: Velocity-based and Bare bones swarm. The evaluation results confirmed that velocity-based outperformed the other one. The second test case compared two PSO variants with different models of influence: best-of-neighborhood and fully-informed swarm. The evaluation results showed that the fully-informed algorithm outperformed the best-of-neighborhood.

Chapter 3

Evolutionary Algorithms

This chapter presents a detailed description of related evolutionary algorithms such as cultural algorithms, particle swarm optimization, and their social variations. In this section, we try to give a brief description of those evolutionary algorithms as a background to our proposed approaches.

3.1 Evolutionary Algorithms

Evolutionary computing is a research area within computer science. As the name suggests, it is a particular flavor of computing, which draws inspiration from the process of natural evolution. In fact, they are computer programs which try to solve and optimize complex problems by simulating the behavior of natural systems. They utilize evolutionary operators such as crossover and mutation to improve the quality of a population of solutions.

Evolutionary algorithms start with a population of randomly generated solutions for a particular problem. Then, the algorithm modifies the solutions in an iteration-based process of some finite number of generations. The modification occurs through applying the so-called evolutionary operators such as crossover and mutation. The evolutionary operators might be binary or unary. For example, the crossover is a

binary operator because it combines two solutions (individuals) to generate a new one. On the other hand, mutation is a unary operator because it makes random modifications on a single solution to improving its performance. The performance of each is evaluated using a fitness function. The fitness function represents the problem that needs to be solved or optimized. Usually, the fitness function is chosen from NP problems which are hard to solve by traditional problem-solving methods. In each iteration of an evolutionary algorithm, the evolutionary operators are applied to the individuals and the best performing offsprings will be transferred to the next generation. This process continues until meeting some stopping criteria which are already defined by a human user. Some examples of evolutionary algorithms are:

1. Genetic Algorithms
2. Evolutionary Programming
3. Differential Evolution

However, the evolutionary algorithms are not restricted to biological processes. Some researchers have drawn inspiration from social systems to find solutions to complex problems. The complex and coordinated behavior of swarms not only fascinates biologists but also is an inspiration to computer scientists. Ant colonies and birds flocking are remarkable examples of coordinated collective behavior that emerges without centralized control. Swarm intelligence is a field of computer science that invents computational methods for solving problems in a way that is inspired by the behavior of real swarms and colonies. Principles of self-organization and communication (local and indirect) are essential to understanding the complex collective behavior. There are different types of swarm-based algorithms such as:

1. Particle Swarm Optimization (PSO)
2. Ant Colony Optimization (ACO)

3. Cultural Algorithms (CA)

3.2 Cultural Algorithms

Cultural Algorithms (CA) were introduced by Reynolds as a type of population-based problem-solving approaches. CA combines biological evolution with socio-cognitive concepts to yield an optimization approach based on a dual inheritance theory. As defined by [Durham], culture is a “system of symbolically encoded conceptual phenomena that are socially and historically transmitted within and between populations”. From the definition, it can be stated that in cultural systems, evolution occurs at two levels: Macro-evolutionary level and micro-evolutionary level. Cultural algorithms define the evolution process through the cooperation of three distinct components:

1. Belief Space (Macro-evolutionary Level)
2. Population Space (Micro-evolutionary Level)
3. Communication Protocol

3.2.1 Belief Space

Belief space keeps different kinds of knowledge obtained from the individuals’ experience during the evolution process. It extracts the knowledge from the population and stores the knowledge in various formats called knowledge sources (KS). Each knowledge source extracts and keeps knowledge about a particular aspect of the search space.

After each iteration, the best performing individuals are chosen and sent to the belief space to be utilized by the knowledge sources. The stored knowledge is used to bias

the search process in the population space. Five types of knowledge sources have been identified in the belief space:

1. Situational Knowledge: Successful exemplars of individuals
2. Normative Knowledge: The best range of value for each parameter
3. Topographic Knowledge: The best areas of the search space
4. Domain knowledge: The domain ranges for all the parameters and the best exemplars
5. Temporal Knowledge: Knowledge about the past changes in a dynamic environment

3.2.2 Population Space

In the population space, any population-based algorithm could be used. In the earlier variants of cultural algorithms, only GA was used. However, researchers utilized other population-based algorithms such as PSO and EP in the population component. In each generation, the best individuals are sent to the belief space to update the knowledge sources. Then, the belief space influences the population through the search process. Figure []

3.2.3 Communication Protocol

In cultural algorithms, the belief space and the population communicate together through a two-way protocol. The population space sends the best individuals to the belief space using the acceptance function to update the knowledge sources. Then, the belief space bias the search direction of the population space through the influence function.

3.3 Social Fabric-based Cultural Algorithms

Social Fabric is a dynamic grid of information flow in which the individuals interactions happen. The fabrics (networks) are created by the connectivity of each agent with other agents in a dynamic structure. The topology of the network controls the rate and type of interactions. Like a multi-population model, there are multiple independent subpopulations (tribes) which are networked together.

In the Social Fabric idea, the influence of the belief space is propagated through the population using a multi-layered network of connections. There are Z tribes comprised of H individuals which form two layers: rudimentary and advanced. The members of each tribe are connected in a regular topology independent from other tribes. At this level, they form the rudimentary layer. From each tribe, the best performing individuals (elites) are connected. These connections create the advanced layer which is responsible for mediating the flow of information between different tribes. Figures 2 and 3 describe the network model.

3.3.1 Evolution Phases

The whole evolution process is split into three phases: seclusion, rapport, and cohesive. In the first stage, each tribe evolves as an independent basic CA model with no communication between tribes. In the second step, the advanced layer is formed by connecting elites of each tribe. Then, tribes start to exchange information during their evolution process. Ultimately, in the third step, all the tribes are merged into one CA model. Then the search process continues until meeting some stopping criteria.

3.3.2 Strategic Neighborhood Restructuring

In Social Fabric literature, Strategic Restructuring is a technique to help individuals to get rid of stagnation where there are copious local optima in the search landscape. Each tribe maintains a particular topology until it gets stagnated in a local optimum for a certain number of iterations. Then, the topology is reinitialized to motivate the individuals for a new search. Here, four types of topologies are utilized: Ring, Mesh, Hybrid-Tree, and Global.

3.3.3 Social fabric based influence function

In this model, the individuals might decide to change their controller KS using the majority of KSs which they receive from their neighborhood. Here, Majority Voting is used to find the controller KS of an individual [?]. In the case of a tie, some tie-breaking rules are deployed such as MFU (most frequently used), LFU (least frequently used), Direct (the direct influencing KS), Random (a random choice), Last-used (the KS which controlled the KS in the last iteration) [?]. In the equation 3.1 the sum denotes the counts of KSs in i th node neighborhood and ψ_i is its controller KS. τ_i is the net affecting KS for the next iteration [?] [?].

$$\tau_i = \sum_{j \in Nbr(i)} m_{ij} + \psi_i \quad (3.1)$$

The Social Fabric approach can be generalized on different topologies as follows:

$$KS(t+1) = \begin{cases} KS_i, & \forall KS_j \in \{KS - KS_i\} \Rightarrow weight(KS_i) > weight(KS_j) \\ KS_{cr}, & otherwise \end{cases} \quad (3.2)$$

Where KS is the set of all knowledge sources, $KS_i, KS_j \in KS$. $weight(KS_i)$ is the number of neighbors that belong to the knowledge source KS_i . KS_{cr} denotes the

knowledge source chosen by a tie-breaking rule.

3.4 Particle Swarm Optimization

The initial idea for particle swarm optimization of Kennedy and Eberhart were essentially aimed at producing computational intelligence by utilizing simple models of social interaction, rather than purely single-agent cognitive capabilities.

In PSO, some simple-structured individuals (the particles) moving around in the search space of a function, and each particle evaluates the objective function based on its current location. Then, each particle determines its movement direction and velocity through the search space by combining its historical best experience and the best (best-fitness) particle in its visible neighborhood in the swarm, with some random fluctuations for keeping diversity. This process repeats at each iteration. Eventually, the swarm as a whole, like a flock of birds collectively searching for food sources, is similar to move around for finding an extremum of the fitness function. Each particle is comprised of a position vector and velocity vector [?]. Particles adjust their velocities and positions as follows:

$$v_{id}^{(t+1)} = w \cdot v_{id}^{(t)} + c_1 r_1 (p_{id} - x_{id}^{(t)}) + c_2 r_2 (p_{gd} - x_{id}^{(t)}) \quad (3.3)$$

$$x_{id}^{(t+1)} = x_{id}^{(t)} + v_{id}^{(t+1)} \quad (3.4)$$

Where v_{id} and x_{id} are the velocity and position of i th particle. c_1 and c_2 are two positive constants. r_1 and r_2 are randomly generated numbers in $[0,1]$ range. w is the inertia weight of which restricts the velocity of a particle. p_{id} is the best experience of the particle and p_{gd} is the best solution in its neighborhood.

3.4.1 Tribe-PSO

Similar to Social Fabric model, the whole population is divided into some tribes. The tribes form a two-layer structure of networks, and the whole process consists of three phases. The individuals in each tribe are completely isolated from other tribes. So, they form the basic layer. On the other hand, the best members of each tribe can see the other tribes and exchange information with their best members as well. The problem-solving process is divided into three phases: isolated, communing, and united. In the first phase, the tribes are completely isolated and there is no exchange of information. In the second phase, the two-layered model is formed and the tribes begin to exchange information through their best performing individuals. In the third phase, all the tribes are merged into a single population. Then, it operates as basic PSO model until meeting some stopping criteria. In fact, the idea of Social Fabric originates from Tribe-PSO. But the network structure is utilized to propagate the influence of knowledge sources in the belief space [?]. In the thesis, we use the social fabric for both PSO interactions and knowledge propagation of the belief space. Figure ?? describes ...

Chapter 4

Proposed Approaches: Improving Robustness in Social Fabric

Inspired by natural systems, many evolutionary algorithms have been proposed to solve various optimization problems. These algorithms are expected to exhibit a consistent problem-solving behavior against different optimization landscapes. Unfortunately, as stated by No Free Lunch (NFL) theorem [?], there is no algorithm better than others over all cost functions. It means, there is no guarantee for an algorithm to work well for all functions if it shows promising results for a particular category of them. Therefore, robustness has been one of the most desired features which motivates researchers to invent new methods which are less dependent on the kind of a problem than others. Here, robustness means we need to develop search strategies which are capable of adapting themselves across different landscapes which might vary regarding aspects such as multimodality and the number of local optima. In this section, we are going to improve the robustness of CAs in both population space and belief space components. In the population space, a new neighborhood restructuring strategy will be proposed which aims at microscopic inspections for stagnation regarding individuals' local experience. Then, in the belief space, a new

model of Normative KS will be proposed, which defines the normative ranges based on Confidence Interval inspired from Inferential Statistics. It stops the normative KS to be affected by sudden fluctuations in the input data.

4.1 Self-organization

A self-organized system is a group of agents with specific behavioral patterns which could not be predicted from the simple behaviors of the individuals who make up the system. Kennedy and Eberhart [1] described self-organization as follows:

”Self-organization The ability of some systems to generate their form without external pressures, either wholly or in part. It can be viewed as a systems constant attempts to organize itself into ever more complex structures, even in the face of the constant forces of dissolution described by the second law of thermodynamics.”

Some of the primary attributes of self-organizing systems are listed below:

1. Self-organizing systems usually exhibit the behaviors that appear to be spontaneous order.
2. Self-organization can be considered as a systems constant attempts to organize itself into a variety of complex structures, even in the face of the forces of deviation (Robustness).
3. The overall status of a self-organizing system is an emergent property of all the ingredients of the system.
4. Interconnected components of the system become organized in a meaningful way based on local interactions of the components.
5. Complex systems have the potential to self-organize themselves.

6. The self-organization process works close to the edge of chaos.

In the section, we will propose a new approach to improve the robustness of cultural algorithms regarding self-organization.

4.2 Irregular Neighborhood Restructuring

As proposed by [?], when the agents of a tribe get stuck in a local optimum, then they choose a topology with fewer connections such as ring topology to motivate exploration. When the agents lack information, the topology is changed to a denser topology like Mesh, and ultimately Global. Whenever a tribe does not make progress for M_{thresh} iterations, then the algorithm decides on upgrading/downgrading the current topology to other topologies. Downgrading happens by moving in the direction of $gbest \rightarrow tree \rightarrow mesh \rightarrow lbest$ and Upgrading occurs in the direction of $lbest \rightarrow mesh \rightarrow tree \rightarrow gbest$. In our proposed approach, the restructuring process occurs at a microscopic level. There is no daemon process for inspecting stagnation. Every agent checks its performance for stagnation. If it gets stuck in a local optimum, it decides to upgrade/downgrade its neighborhood. The proposed strategy is described in the algorithm ?? . If the particle is the best in its neighborhood, it starts to decrease the number of connections to motivate exploration. However, if it is not the best, it needs to increase the neighborhood size to facilitate exploitation. Here, the topology is dynamic and irregular and the neighborhood size is different for each agent. So, it could be considered as a heterogeneous variant of CAs [?]. The ultimate topology of the fabric is formed through local interactions of the agents. In this way, it improves the robustness of CAs regarding self-organization concept [?].

4.3 Confidence based Normative Knowledge Source

The current version of Normative KS is quite vulnerable to temporary fluctuations in the pattern of input data. The size of Normative ranges is subject to change dramatically with any temporal fluctuations. In this section, we are going to replace the ranges in the standard Normative KS with Confidence Interval from Inferential Statistics. Confidence-based changes are more robust against instantaneous changes in the input pattern. The update process of standard normative ranges is as follows:

$$lb_j(t+1) = \begin{cases} x_{i,j}(t), & x_{i,j}(t) \leq lb_j(t) \text{ or } f(X_i) < PL_j(t) \\ lb_j, & \text{otherwise} \end{cases} \quad (4.1)$$

$$ub_j(t+1) = \begin{cases} x_{i,j}(t), & x_{i,j}(t) \geq ub_j(t) \text{ or } f(X_i) < PU_j(t) \\ ub_j, & \text{otherwise} \end{cases} \quad (4.2)$$

The following is the definition of normative ranges based on Confidence Interval concept [?].

$$lb_j = \bar{x}_j - q \cdot \frac{\sigma_j}{\sqrt{n}} \quad (4.3)$$

$$ub_j = \bar{x}_j + q \cdot \frac{\sigma_j}{\sqrt{n}} \quad (4.4)$$

where \bar{x}_j is the mean of j th dimension, σ_j is the standard deviation of j th dimension and q is the confidence coefficient

Chapter 5

Experimental Setup

In this chapter, we explain the experimental setup, how to set the parameters and the optimization functions as the benchmark to evaluate the efficiency of our proposed approaches.

5.1 Description

In this section, we describe IEEE-CEC2015 as the function set which we have chosen for our benchmark optimization purpose. IEEE-CEC2015 is a set of 15 functions with different properties such as multi-modality, copious local optima, and non-separability. All test functions are minimization problems defined as follows:

$$Y = f(x_1, x_2, x_3, \dots, x_D) \quad (5.1)$$

where D is the number of dimensions.

Before evaluation, all the vectors are shifted and rotated as described in [1]. $\mathbf{o}_i = [o_{i1}, o_{i2}, \dots, o_{iD}]$ is the shifted global optimum, which is randomly distributed in $[-80, 80]$. For convenience, the same search ranges are defined for all test functions as $[-100, 100]^D$.

5.2 Benchmark Functions

In this section, we describe the IEEE-CEC2015 functions and their properties as a well-known benchmark adopted by many researchers to evaluate innovative ideas in solving complex optimization problems.

5.2.1 Unimodal Functions

Rotated Bent Cigar Function

$$f_1(x) = x_1^2 + 10^6 \sum_{i=2}^D x_i^2 \quad (5.2)$$

$$F_1(x) = f_1(M(x - o_1)) + F_1^* \quad (5.3)$$

Properties:

1. Unimodal
2. Non-separable
3. Smooth but narrow ridge

Rotated Discus Function

$$f_2(x) = 10^6 x_1^2 + \sum_{i=2}^D x_i^2 \quad (5.4)$$

$$F_2(x) = f_2(M(x - o_2)) + F_2^* \quad (5.5)$$

Properties:

1. Unimodal
2. Non-separable
3. With one sensitive direction

5.2.2 Simple Multimodal Functions

Shifted and Rotated Weierstrass Function

$$f_3(x) = \sum_{i=1}^D \left(\sum_{k=0}^{kmax} [a^k \cos(2\pi b^k (x_i + 0.5))] \right) - D \sum_{k=0}^{kmax} [a^k \cos(2\pi b^k \cdot 0.5)] \quad (5.6)$$

where a=0.5, b=3, and kmax=20.

$$F_3(x) = f_3\left(M\left(\frac{0.5(x - o_3)}{100}\right)\right) + F_3^* \quad (5.7)$$

Properties:

1. Multi-modal
2. Non-separable
3. Continuous but differentiable only on a set of points

Shifted and Rotated Schwefel's Function

$$f_4(x) = 418.9829 \times D - \sum_{i=1}^D g(z_i), z_i = x_i + 4.209687462275036e + 002 \quad (5.8)$$

$$F_4(x) = f_4\left(M\left(\frac{1000(x - o_4)}{100}\right)\right) + F_4^* \quad (5.9)$$

Properties:

1. Multi-modal
2. Non-separable
3. Local optima's number is huge and second better local optimum is far from the global optimum.

Shifted and Rotated Katsuura Function

$$f_5(x) = \frac{10}{D_2} \prod_{i=1}^D \left(1 + i \sum_{j=1}^{32} \frac{|2^j x_i - \text{rand}(2^j x_i)|}{2^j}\right)^{\frac{10}{D^{1.2}}} - \frac{10}{D^2} \quad (5.10)$$

$$F_5(x) = f_5(M(\frac{5(x - o_5)}{100})) + F_5^* \quad (5.11)$$

Properties:

1. Multi-modal
2. Non-separable
3. Continuous everywhere yet differentiable nowhere

Shifted and Rotated HappyCat Function

$$f_6 = |\sum_{i=1}^D x_i^2 - D|^{0.25} + \frac{0.5 \sum_{i=1}^D x_i^2 + \sum_{i=1}^D x_i}{D + 0.5} \quad (5.12)$$

$$F_6(x) = f_6(M(\frac{5(x - o_6)}{100})) + F_6^* \quad (5.13)$$

Properties:

1. Multi-modal
2. Non-separable

Shifted and Rotated HGBat Function

$$f_7 = |(\sum_{i=1}^D x_i^2)^2 - (\sum_{i=1}^D x_i)^2|^{0.25} + \frac{0.5 \sum_{i=1}^D x_i^2 + \sum_{i=1}^D x_i}{D + 0.5} \quad (5.14)$$

$$F_7(x) = f_7(M(\frac{5(x - o_7)}{100})) + F_7^* \quad (5.15)$$

Properties:

1. Multi-modal
2. Non-separable

Shifted and Rotated Expanded Griewank's plus Rosenbrock's Function

$$f_8(x) = f_{11}(f_{10}(x_{x_1, x_2})) + f_{11}(f_{10}(x_{x_2, x_3})) + \cdots + f_{11}(f_{10}(x_{x_{D-1}, x_D})) + f_{11}(f_{10}(x_{x_D, x_1})) \quad (5.16)$$

$$F_8(x) = f_8(M(\frac{5(x - o_8)}{100}) + 1) + F_8^* \quad (5.17)$$

Properties:

1. Multi-modal
2. Non-separable

Shifted and Rotated Expanded Scaffer's F6 Function

$$g(x, y) = 0.5 + \frac{(\sin^2(\sqrt{x^2 + y^2}) - 0.5)}{(1 + 0.001(x^2 + y^2))^2} \quad (5.18)$$

$$f_9(x) = g(x_1, x_2) + g(x_2, x_3) + \cdots + g(x_{D-1}, x_D) + g(x_D, x_1) \quad (5.19)$$

$$F_9(x) = f_9(M(x - o_9) + 1) + F_9^* \quad (5.20)$$

Properties:

1. Multi-modal
2. Non-separable

5.2.3 Hybrid Functions

Hybrid Function 1

This function is a hybrid of three functions: Modified Schwefel's function, Rastrigin's function, and High Conditioned Elliptic function.

$$F_{10}(x) = 0.3 \times F_4(x) + 0.3 \times F_{12}(x) + 0.4 \times F_{13}(x) \quad (5.21)$$

Hybrid Function 2

This function is a hybrid of 4 functions: Griewank's function, Weierstrass function, Rosenbrock's function, and Scaffer's F6 function.

$$F_{11}(x) = 0.2 \times F_{11}(x) + 0.2 \times F_3(x) + 0.3 \times F_{10}(x) + 0.3 \times F_9(x) \quad (5.22)$$

Hybrid Function 3

This function is a hybrid of 5 functions: Katsuura function, HappyCat function, Expanded Griewank's plus Rosenbrock's function, Modified Schwefel's function, and Ackley's function.

$$F_{12}(x) = 0.1 \times F_5(x) + 0.2 \times F_6(x) + 0.2 \times F_8(x) + 0.2 \times F_4(x) + 0.3 \times F_{14}(x) \quad (5.23)$$

5.2.4 Composite Functions

The general format of composite function is as follows:

$$F(x) = \sum_{i=1}^N \{\omega_i \times [\lambda_i g_i(x) + bias_i]\} + f^* \quad (5.24)$$

$F(x)$: composition function

$g_i(x)$: i^{th} basic function used to construct the composition function

N : number of basic functions

ω_i : new shifted optimum position for each $g_i(x)$, define the global and local optima's position.

$bias_i$: defines which optimum is global optimum.

σ_i : used to control each $g_i(x)$'s coverage range, a small σ_i give a narrow range for that $g_i(x)$

λ_i : used to control each $g_i(x)$'s height

w_i : weight value for each $g_i(x)$, calculated as follows:

$$w_i = \frac{1}{\sqrt{\sum_{j=1}^D (x_j - o_{ij})^2}} \exp\left(-\frac{\sum_{j=1}^D (x_j - o_{ij})^2}{2D\sigma_i^2}\right) \quad (5.25)$$

Now, the value of ω_i weight is calculated as:

$$\omega_i = \frac{w_i}{\sum_{i=1}^n w_i} \quad (5.26)$$

Composition Function 1

Five types of basic functions are combined to construct this function:

$N=5$, $\sigma = [10, 20, 30, 40, 50]$, $\lambda = [1, 1e-6, 1e-26, 1e-6, 1e-6]$, $bias = [0, 100, 200, 300, 400]$

g_1 : Rotated Rosenbrock's Function f_{10}

g_2 : High Conditioned Elliptic Function f_{13}

g_3 : Rotated Bent Cigar Function f_1

g_4 : Rotated Discus Function f_2

g_5 : High Conditioned Elliptic Function f_{13}

Properties:

1. Multi-modal
2. Non-separable
3. Asymmetrical
4. Different properties around different local optima

Composition Function 2

Three types of basic functions are combined to construct this function:

$N=3$, $\sigma = [10, 30, 50]$, $\lambda = [0.25, 1, 1e-7]$, $bias = [0, 100, 200]$

g_1 : Rotated Schwefel's function f_4

g_2 : Rotated Rastrigin's function f_{12}

g_3 : Rotated High Conditioned Elliptic function f_{13}

Properties:

1. Multi-modal
2. Non-separable
3. Asymmetrical
4. Different properties around different local optima

Composition Function 3

Five types of basic functions are combined to construct this function:

$N = 5$, $\sigma = [10, 10, 10, 20, 20]$, $\lambda = [10, 10, 2.5, 25, 1e - 6]$, $bias = [0, 100, 200, 300, 400]$

g_1 : Rotated HGBat Function f_7

g_2 : Rotated Rastrigin's Function f_{12}

g_3 : Rotated Schwefel's Function f_4

g_4 : Rotated Weierstrass Function f_3

g_5 : Rotated High Conditioned Elliptic Function f_{13}

Properties:

1. Multi-modal
2. Non-separable
3. Asymmetrical
4. Different properties around different local optima

5.3 Parameters Settings

We used the same initial values for the parameters as presented in []:

- Population Size: 90
- Number of tribes: 10
- The window-size for applying the social influence with restructuring (Wsize): 20
- The proportion of the best agents to be sent to the belief space (Pa): 0.25
- The number of elites in each tribe (nelite): 2
- The threshold trigger for restructuring (Mthresh): 50
- Tie-breaking rule (TieR): MFU
- The maximum number of function evaluations (MaxFES): 150000
- The social factor (Sf) is an integer value variable that takes a value between 0 and 3. Note that these figures correspond to topologies, lbest, von-Neumann, n-star-bus, and gbest.

Here, we use MaxFES as a stopping criterion. There are 20 independent runs for each function. And each function is tested for 10 and 30 number of dimensions.

Appendix A

Additional Information

This is a good place to put tables, lots of results, perhaps all the data compiled in the experiments. By avoiding putting all the results inside the chapters themselves, the whole thing may become much more readable and the various tables can be linked to appropriately.

The main purpose of an Appendix however should be to take care of the future readers and researchers. This implies listing all the housekeeping facts needed to continue the research. For example: where is the raw data stored? where is the software used? which version of which operating system or library or experimental equipment was used and where can it be accessed again?

Ask yourself: if you were given this thesis to read with the goal that you will be expanding the research presented here, what would you like to have as housekeeping information and what do you need? Be kind to the future graduate students and to your supervisor who will be the one stuck in the middle trying to find where all the stuff was left!

Bibliography

- [1] Cliff Atkinson. *Beyond Bullet Points: Using Microsoft Office PowerPoint 2007 to Create Presentations That Inform*. Microsoft Press, 2008.
- [2] Wayne Booth. *The Craft of Research*. University of Chicago Press, 2003.