Accepted Manuscript

The Evolution of System Reliability Optimization

David W. Coit , Enrico Zio

Please cite this article as: David W. Coit , Enrico Zio , The Evolution of System Reliability Optimization, *Reliability Engineering and System Safety* (2018), doi: https://doi.org/10.1016/j.ress.2018.09.008

Highlights

- Research pertaining to system reliability optimization has been categorized chronologically into three eras.

- The first era, the era of mathematical programming involved rigorous optimization methods, yet the problems are not always realistic.

- The second era, the era of pragmatism, involved expanding the problem domain to include a broader range of problems, and more realistic problems.

- The final and current era, the era of active reliability improvement involves dynamic reliability optimization models responding to changing conditions and data.

- System reliability optimization problems remain challenging, but important, while both the problems and corresponding solution methods evolve.

The Evolution of System Reliability Optimization

David W. Coit[a], Enrico Zio[b,c,d]

[a]Department of Industrial & Systems Engineering, Rutgers Uniuversity, Piscataway, NJ, USA
[b]Chair on System Science and the Energetic Challenge, EDF Foundation, Laboratoire Genie Industriel, CentraleSupélec, Université Paris-Saclay, France
[c]Energy Department, Politecnico di Milano, Italy
[d]Department of Nuclear Engineering, College of Engineering, Kyung Hee University, Republic of Korea

Abstract – System reliability optimization is a living problem, with solutions methodologies that have evolved with the advancements of mathematics, development of new engineering technology, and changes in management perspectives. In this paper, we consider the different types of system reliability optimization problems, including as examples, the redundancy allocation problem (RAP), the reliability allocation problem and the reliability-redundancy allocation problem (RRAP), and provide a flow of discussion and analysis on the evolution of the approaches for their solutions. We consider the development and advancement in the fields of operations research and optimization theory, which have allowed the formalization and continuous improvement of the methods and techniques to address reliability design problems of even very complex systems in different technological domains. Technological advances have naturally brought changes of perspectives in response to the needs, interests and priorities of the practical engineering world. The flow is organized in a structure of successive "Eras of Evolution," namely the Era of Mathematical Programming, the Era of Pragmatism, the Era of Active Reliability Improvement. Insights, challenges and opportunities are highlighted.

1. Introduction

"Success is walking from failure to failure with no loss of enthusiasm" – Winston Churchill

Reliability engineering is a formal engineering discipline, founded on mathematical principles, particularly of probability theory and statistics, for systematically and rigorously analyzing functional problems in components and systems with the aim to produce a reliable design. As an engineering discipline, reliability aims at analyzing and evaluating the ability of products and services to perform the functions that they are intended to provide by design.

While technology improves and advances, the complexity of modern engineered systems also increases. At the same time, consumers' expectations for high functionality, high performance and high reliability increase, leading to challenges and opportunities. Then, although system reliability optimization problems have been studied, analyzed, dissected and

reanalyzed, the continuous rise of new challenging problems demonstrates that this general research area will never be devoid of interesting problems to be solved.

On one side, the development and advancement in the fields of operations research and optimization theory have allowed the formalization and continuous improvement of the methods and techniques to address reliability design problems of even very complex systems in different technological domains. On another side, the evolution of technology, the advancement of research ideas and theories, have naturally brought changes of perspectives, in response to the needs, interests and priorities of the developing practical engineering world. So, the development and application of formal optimization methods to the practical goal of achieving maximum reliability under various physical and economic constraints, has remained an ongoing topic of scientific development.

In formal terms, the task of optimization involves formally conceptualizing the decision variables, the constraints and the single or multiple objective functions that describe the performance of the engineering design problem, and searching for the combination of values of the decision variables that achieve the desired goals with respect to the objective functions. Whether expressed explicitly in mathematical terms or not, every engineering design problem has design objectives which should be maximized or minimized or designed to achieve some acceptable requirement. When there is a single predominant objective that can be expressed with a series of explicit mathematical equations, then the problem can potentially be solved using mathematical programming methods or useful heuristics. Researchers working within the broader mathematical programming community are continually developing new methods and algorithms to solve broader classes of problems, larger and more difficult problems, and to do so more efficiently than before.

1.1 Eras of Research Evolution

The research in complex system reliability optimization has evolved as a continuum of ideas and techniques. This evolution can be loosely and chronologically classified into the following three eras:

- Era of Mathematical Programming
- Era of Pragmatism
- Era of Active Reliability Improvement

The era of mathematical programming is preceded by the original development of innovative groundbreaking methods, such as dynamic programming and the simplex algorithm (or linear programming), for which the reliability optimization problem has served as a very interesting and practical vehicle to demonstrate the methods and apply them. Yet, practicing reliability analysts recognized the limitations of these methods in practical applications, mostly because only problems that could be formulated to strictly meet the assumptions required by the methods could be solved, and this is rarely the case in practice. Furthermore, only small and/or unrealistic problems could be solved, because of computation limitations at the time.

Driven by the desire to apply reliability optimization in practice, the era of pragmatism evolved and became increasingly important. New problems were solved and new methods developed in response to the pressing needs to consider and integrate into the problem some critical issues that could not be readily accommodated by the rigorous mathematical methods. For example, actual engineering systems problems could fall outside the assumptions required by the methods. Analysts were interested in complex forms of redundancy perhaps mixing functionally equivalent, but different components, whose failure behavior may not be simply described as a transition from one functioning state to a failure state, but rather as a process of transition across multiple states. To address a broader range of problems, compromises could be accepted, thus expanding the practical usefulness and applicability of the optimization methods.

In these previous eras, reliability optimization had been mostly considered as singular, static analyses to be conducted and implemented to yield the desired design reliability of the system. Under this view, common assumptions were made on the existence of populations of homogeneous components and systems, sharing the same failure behavior but with failure occurrences being independent from one another. Furthermore, the implied assumption was that the conditions defined or considered when conducting the analysis are static and remain unchanged throughout the horizon of the analysis (the often-called mission time).

When changes occur during the system lifetime, then the results of the analysis are simply no longer valid and applicable. We are currently experiencing another industrial revolution, particularly driven by the increase in information sharing, data availability and computational capabilities. In particular, with the proliferation of sensors, environmental stresses, usage stresses, failure data, etc., can be collected and processed at regular time intervals,

and the advancements in information knowledge that these can bring on the states of the systems, offer new opportunities of development for the analysis and assessment of reliability. The era of active reliability improvement is, then, ongoing and it is visionary in recognizing that components and system conditions change throughout their lives, and system reliability optimization methods, to be practically useful, need to dynamically respond to these changes.

Within each of the three macro-eras discussed above, further sub-classification can be introduced, in a specific and unique way. In the era of mathematical programming, further sub-classification can be done based on the specific mathematical algorithms developed to solve the reliability design problem. For the era of pragmatism, sub-classification can refer to which practical consideration in the design problem, which previously could not be analyzed, was now readily being addressed by the available optimization models. Within the era of active reliability improvement, sub-classification can be based with reference to the available data and new models for real reliability improvement.

2. Reliability Optimization Problems

System reliability design problems have multiple, and often competing objectives. However, there are some universal ones, including reliability (to be maximized) and cost (to be minimized). Often, the approach taken is to follow a prioritization of the objectives by the decision-makers, select the most important objective as the objective function and constrain the other objectives within acceptable limits.

For each formulation to be studied or solved, system reliability optimization problems must have three elements: decision variables, constraints and an objective function or functions. The decision variables are those variables that can be changed or decisions that can be made to improve performance, with respect to the objective function or functions. Examples of decision variables include component type (with its intrinsic characteristic of failure behavior and reliability), redundancy configuration in the system and others. Constraints are mathematical expressions of practical limitations, such as monetary budget or acceptable reliability, which limit the choice of decision variables in relation to their feasibility of respecting the constraints. The objective function measures the performance of the system for given values of the decision variables, and thus, enables the decision on the optimal combination of variables values for the optimal solution. The objective function can often be the system reliability to be maximized, or

the system cost to be minimized.

Different forms exist of the system reliability optimization problem. Three typical ones are the redundancy allocation problem (RAP), the reliability allocation problem and the reliability-redundancy allocation problem (RRAP). The solution methods for each problem are obviously different, because of the assumptions and structure of the problem. A thorough review of system reliability optimization is given by Kuo et al [1-3].

2.1 Redundancy Allocation Problem

The most widely studied reliability optimization problem is the RAP. For many systems composed of discrete component types, with fixed cost, reliability and weight, in mathematical terms system design becomes a combinatorial optimization problem. For providing the demanded system functions, there may often be alternative component types available, at different cost, reliability, weight and other properties. The practical problem is to select the optimal combination of components types (decision variables) to collectively meet reliability, weight, etc. (constraints) at a minimum cost (objective function), or alternatively, to maximize reliability (objective function) while achieving given minimum (or maximum) values of other system properties (constraints).

For the mathematical formulation of this problem, we can consider that there are m_i discrete component type choices available for each subsystem (i=1, ..., s), which the system is formed of. Figure 1 shows a typical example of a system with a number s of k-out-of-n logic subsystems. If k is equal to 1 for each subsystem, then this is a simple series-parallel system. For each subsystem, k_i components must be selected from the m_i available choices (e.g., assuming an unlimited amount available for each of the components). The optimal decision is driven by the possibility of placing additional components in parallel in any of the subsystems and/or adding levels of redundancy ($> k_i$) of lower reliability components as an alternative to using more reliable (and expensive) ones. In other words, there is a large number of possible solutions to test, even for relatively small problems (small number of subsystems, small number of components, small number of components types, etc.).

The RAP for a series-parallel system, as the one shown in Figure 1, can be formulated as to maximize reliability or minimize cost, under various constraints. Often the RAP is solved for series-parallel systems, but other system structures have been considered, as well as more

advanced forms of series-parallel systems, including those with several failure modes, phase mission types, systems with uncovered failures and systems with imperfect fault coverage.

RAP has proven to be quite difficult to solve. Chern [4] showed that the problem is NP-hard and many different optimization approaches have been used to determine optimal or "very good" (i.e., near-optimal) solutions, including (1) dynamic programming, (2) integer programming, (3) mixed integer and nonlinear programming, or (4) evolutionary algorithms.

2.2 Reliability Allocation Problem

The reliability allocation problem and RAP are fundamentally different problems. For the reliability allocation problem, the system structure is fixed and the component reliability values are continuous decision variables. For this problem, there is no general restriction on the system structure. An example of a general system are depicted in Figure 2. Component cost and other parameters are defined as mathematical functions of the component reliability. Increasing the component reliability (and thus, the system reliability) increases the cost, weight and other factors, which may be included as part of the constraints or the objective function. Here, also, the goal of the optimization is typically to maximize system reliability or minimize system cost, and since the decision variables are continuous, different forms of non-linear programming can be used to determine the optimal solutions. To assure that the constraints are satisfied, Lagrangian multipliers are often introduced as part of the objective function.
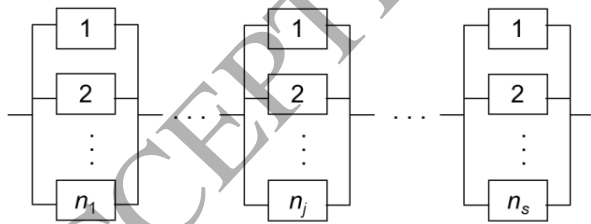
Figure 1: Series-Parallel System   Figure 2: General System Structure
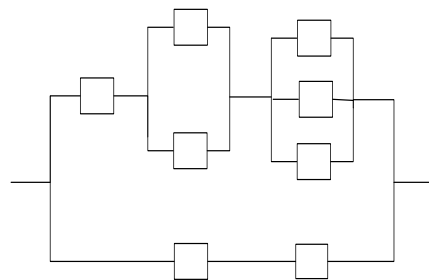
2.3 Reliability-Redundancy Allocation Problem (RRAP)

The reliability-redundancy allocation problem is the most general problem formulation. The system is composed of one or more "subsystems," i.e., collections of logically connected sets of components. Each subsystem has x_i components with reliability of r_i as decision variables. The problem is then to optimally allocate redundancy and reliability to the components of each

subsystem with an objective to maximize the overall system reliability. The system and subsystems can be in any arrangement, as depicted in Figure 2. Again, typically the objective of the optimization is to maximize system reliability or minimize system cost.

RAP is often considered for series-parallel systems, but other system structures can be considered as well. The reliability allocation problem and RRAP has been applied to many different system structures, including common structures (series, parallel, etc.), but also consecutively connected systems, sliding window systems, flow transmission systems, common bus systems and others.

2.4 Component Assignment and Sequencing Problems

Two other related system reliability problems are assignment or sequencing of components within a system. For assignment problems, there is typically a defined system structure, but the available components are assigned to specific locations with the system. Sequencing is particularly interesting and important for systems with standby redundancy, where the components are activated as needed in accordance with a defined sequence. Sequencing problems are solved for optimization for systems with consecutively connected systems and/or standby components with cold or warm redundancy required a defined activation sequence.

The problem of assignment of components to positions within a system to maximize reliability was originally presented by Derman et al [5, 6]. An important early research effort [7] defined optimal assignments for different system structures. The related problem involves sequencing of redundant components. More recently, sequencing of standby redundant components has been considered by researchers to maximize system reliability [8, 9].

2.5 Other Optimization Problems

There have been many other related system reliability optimization problems as well. There have been extensions of the original problems, as well as related problems including spares provisioning, optimization of series-parallel topology, optimal load distribution, optimal mission abort policy, test plan allocation, preventive maintenance optimization and others [2].

3. Era of Mathematical Programming

Methods initially developed to solve complex system reliability optimization problems can be referred to as belonging to the era of mathematical programming. The emphasis was on

applying advanced mathematics to obtain mathematically provable optimal solutions. The priority was on solving problems to optimality, but in doing so, often the structure and size of the problems were limited to be able to apply the rigorous mathematical model. The problems being solved in the end were rarely realistic or indicative of actual design problems. Assumptions were often introduced for mathematical convenience, and problems that did not meet those conditions were avoided. However, these original methods are very important and influential, and they have served as the foundation for much of the research and development work in quantitative reliability engineering that followed.

Solution of complex system optimization problems was possible because of advancements in operations research theory and the development of new algorithms. A key mathematical challenge was to find an efficient way to at least approximate solutions to problems otherwise unsolvable with classic analytical methods. Numerical and mathematical approaches were introduced to successfully solve such problems, with the turning point having been the realization of the first computer, which provided the possibility to perform sets of operations for handling large numbers of data in a time much shorter than previously possible.

The newborn field of computer science sparked the mind of several mathematicians that tried to formulate methods to use computers to help solving practical problems involving high computational efforts. The field of mathematical optimization was then born. Several pioneers from the first half of the 20th century contributed to develop formulations and algorithms to be implemented in computing machines to solve difficult optimization problems, including from the fathers of linear programming, Leonid Kantorovich and George Dantzig, Richard Bellman, originator of dynamic programming, to the founders of evolutionary algorithms, such as Lawrence Fogel, John Holland, Ingo Rechenberg and Hans-Paul Schwefel. Their ideas were the milestones from which other variants of the methods developed until today.

When mathematical programming methods associated with the field of operations research were being developed and formalized, but still in their infancy, researchers were searching for interesting applications. Maximization of system reliability was considered an attractive application. Indeed, it is a difficult and challenging problem, yet practical and useful to demonstrate the utility of new mathematical programming algorithms. Typical formulations of the problem are challenging, with a highly nonlinear objective function and often integer

decision variables.

Dynamic programming was applied to the system reliability problem as one of the very first applications explored. Considering the initial formulations of RAP, dynamic programming could almost be directly applied to efficiently obtain optimal solutions. The problem was that it was difficult to extend it to solve more realistic or actual design problems. Linear programming, or the simplex algorithm, is a very powerful advancement, which allowed for very large linear problems to be solved. However, system reliability is a highly nonlinear objective function, so researchers had to be creative to exploit the power of the simplex algorithm to solve reliability problems. RRAP is nonlinear with both continuous and integer decision variables creating another very challenging problem, that was often solved using some variation of nonlinear programming.

3.1 Dynamic Programming

Dynamic programming was originated in 1954 by Richard Bellman [10], and system reliability optimization was among the first problems studied. The aim was to treat mathematical problems arising from the study of multi-stage decision processes. The key advancement, as compared to previous methods, relies on the fact that when analyzing these problems, not all the possible sequences of the present and following stages are needed, i.e., it is possible, instead, to state general conditions to determine for each stage the most suitable decision according to the current state only, whereas classical approaches gathered information about all the sequences, making the calculation cumbersome and unpractical [10].

A problem solvable by dynamic programming can be performed as a system, described by a set of quantities, the state parameters, that undergo a state variation caused at a certain time t by a decision made by the user. The solution aims at taking an initial decision for guiding the future ones so that it is possible to maximize a given objective function of the state parameters. In many cases the number of parameters considered to make the decision is very large, especially when considering stochastic processes in which decisions determine a distribution of outcome states. In these cases, the approach allows reducing the dimension of the problem by focusing on the current time. To perform the dynamic programming optimization, the state parameters and the sequence of decisions to analyze, i.e., a policy, are needed. The optimal policy is, then, the one that determines the decision required at each time with respect to the current state of the

system.

Bellman [11] and Bellman and Dreyfuss [12, 13] demonstrated that an optimal solution to the RAP could be found using dynamic programming. In their problem, there was only one component choice for each subsystem, and the objective was to maximize reliability with a single cost constraint. For each subsystem, the problem was to identify the optimal levels of redundancy. A well-known disadvantage of dynamic programming formulations is the difficulty of efficiently solving problems with multiple constraints.

Fyffe, Hines and Lee [14] used a dynamic programming approach to solve a more difficult design problem. Their problem involved a system with 14 subsystems and cost and weight constraints. For each subsystem in the Fyffe formulation, there are three or four different component choices each with different reliability, cost and weight. However, several of these component choices are dominated by other competing choices. Similar to Bellman, only 1-out-of-n redundancy was considered. To accommodate multiple constraints within a dynamic programming formulation, they used a Lagrangian multiplier for the weight constraint within the objective function.

Instead of using Lagrangian multipliers, Nakagawa and Miyazaki [15] used a surrogate constraint combining the cost and weight constraints into one. They then solved a series of problem iterations with different surrogate multipliers, with a heuristic to successively update the surrogate multipliers. Stopping criteria was provided to identify cases when their algorithm would not lead to an optimal solution. The algorithm was demonstrated by solving 33 variations of the Fyffe problem with different weight constraints. Of the 33 problems, they found optimal solutions to 30 of the problems. Otherwise, the final solution was not feasible (although there are feasible solutions to the problem).

3.2 Linear Programming/Integer Programming

Linear programming (LP) and Integer Programming (IP) are powerful methods to find the maximum or minimum of a linear function describing the performance under assessment, which is called the objective function. A standard mathematical definition is the following:

$$\max \{\mathbf{cx}; A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\} \text{ or } \min \{\mathbf{cx}; A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$$

For this formulation $y = \mathbf{cx}$ is the objective function to be maximized or minimized, $\mathbf{x}$ is

the vector of non-negative variables to be found, **b** and **c** are vectors of known coefficients, and A is a matrix of known coefficients that when multiplied by **x** have to satisfy the constraints expressed by the vector of coefficients **b**. The problem is defined within a convex polyhedron-shaped feasible region, intersection of finitely half-spaces represented by linear equalities/inequalities.

This optimization framework was initially used by the Soviet economist Leonid Kantorovich who was trying to organize the actions of soldiers to decrease expedition costs and increase enemy losses. At the same time another economist, T.C. Koopmans was working on the applicability of linear programming to solve classical problems. Their work was recognized by the Nobel prize in economics in 1975. Following that, mathematician George Dantzig developed an LP methodology to solve optimization problems, providing a formal proof of the solution [16]. One of the most important achievements was the reduction of the possible solutions, and therefore, the advantage of the method in terms of computing power needed.

If the objective is to maximize reliability or minimize cost given multiple nonlinear but separable constraints, many variations of the problem can be transformed into an equivalent integer programming problem using 0-1 decision variables. This was originally demonstrated by Ghare and Taylor [17] who used a branch-and-bound approach to solve many randomly generated RAPs with 30 subsystems with 15 constraints, and 99 subsystems with 10 constraints. Ghare and Taylor assumed that there was only one component choice for each subsystem and redundancy was active always 1-out-of-n redundancy.

Bulfin and Liu [18] also used an IP approach to solve the RAP. They developed one heuristic and two exact algorithms to be applied depending on the problem structure. They formulated the problem as a knapsack problem and used a surrogate constraints approach, similar to Nakagawa and Miyazaki [15]. The surrogate multipliers were approximated as the optimal Lagrangian multipliers as found by subgradient optimization. Bulfin and Liu formulated the Fyffe problem and its variations as integer programs and solved the 33 problems previously investigated by Nakagawa and Miyazaki, and other examples as well. They also considered only subsystems with 1-out-of-n redundancy. Coit and Liu [19] extended their aproach to k-out-of-n redundancy subsystems if no mixing of component types is allowed within the subsystem parallel structure.

Other examples of IP solutions to the redundancy allocation problem were presented by Misra and Sharma [20], Gen, Ida, Tsujimura and Kim [21], and Gen, Ida and Lee [22]. Misra and Sharma presented a very fast and useful algorithm to solve integer programming problems formulated like those of Ghare and Taylor [17]. Gen, Ida, Tsujimura and Kim and Gen, Ida and Lee formulated the problem as a multi-objective decision making problem with distinct goals for reliability, cost and weight.

For other system reliability applications, LP has been proved useful in the context of structural system reliability by Corotis & Nafday [23]. They used LP to identify the most critical failure mode for a structural system. More recent papers [24] demonstrate that LP is particularly useful in structural system reliability analysis. The LP bounds can be applied for any type of system and for different component probabilities. These bounds are the narrowest possible bounds that one can be obtained for a system, for any specified information for the component failure probabilities.

The main drawback of using LP or IP is that the size of the problem to be solved, and its computational cost, increases exponentially with the number of components, questioning its efficiency when it comes to realistic, complex systems. An approach has been proposed to overcome this issue and extend the applicability of LP. Decomposing the entire system into subsystems based on failure modes can be applied to identify component state probabilities and joint probabilities of the states of a small number of components. It can also provide bounds for the failure probability of large systems. This is particularly useful when other methods are not applicable. This approach has been presented by Chang & Mori [25]. The idea is the development of a relaxed LP (RLP) bounds method to reduce the number of design variables using the universal generating function (UGF) [26].

RLP bounds method can be applied to a single series or parallel system, but it is not applicable to a general system that consists of both series subsystems and parallel subsystems. For this reason, an additional assumption can be made to obtain the Strategic Relaxed Linear Programming (SRLP). After decomposing the system according to different failure modes, each critical failure mode is also considered as a system (or subsystem) itself. The bounds on the system failure probability can be computed by the RLP bounds method if it is a series or parallel system, and the bounds on its joint failure probability can also be computed by the RLP bounds

method. The bounds estimated by the RLP bounds method are, then, used as constraints in solving the LP problem for estimating the failure probability of the entire system.

3.3 Nonlinear Programming

Nonlinear programming (NLP) refers to a collection of optimization methods defined by the same main principles of linear programming, with the difference that the objective function and/or its constraints, and therefore the feasible region of the problem, are defined with at least one nonlinear equation. The addition of nonlinear equations makes the optimization problem much more difficult to be solved, for example:

- In a nonlinear function it is hard to assess whether a maximum is local or global, and unlike linear functions where a max/min location is restricted to the borders of the feasible region, for a nonlinear function it can be in the interior of the feasible region.

- If the objective or any constraints are non-convex, the problem may have multiple disconnected feasible regions and multiple locally optimal points within such regions.

- The numerical method chosen to get to the solution may cause two different starting points to lead to two different solutions.

- It is difficult to ensure that the constraints applied to the problem meet the requirements of the feasible region.

- A tolerance region for the solution has to be considered with a proper uncertainty.

NLP solvers generally attempt to solve the problem by computing gradient values at various trial solutions, and moving in the direction of the negative gradient (when minimizing, positive gradient when maximizing). They usually also exploit second derivative information to follow the curvature as well as the direction of the problem functions. To solve constrained problems, NLP solvers must take into account feasibility and the direction and curvature of the constraints as well as the objective. A review of nonlinear programming optimization methods is provided by Floudas [27].

Mixed integer and nonlinear programming have been effectively used to solve the redundancy allocation problem. Considering reliability optimization, important research contributions were provided by Tillman, Hwang and Kuo [28, 29]. In these problems, component reliability is a continuous decision variable and component cost is expressed as a function of reliability and other parameters.

3.4 Evolutionary Algorithms

Evolutionary algorithms (EA) are a group of optimization methods that perform their task with a built-in ability to evolve. EA have the three following common features:

1. Population-based, i.e., they handle a group of solutions, the population, manipulated in different ways to optimize the problem;

2. Fitness-oriented, meaning that EAs favor individuals (a solution belonging to a population) which are fitter than others according to established criteria. Each individual has a gene representation which is its code together with a performance evaluation, i.e., its fitness value. Choosing fitter individuals drives the optimization and the convergence of the algorithm;

3. Variation-driven: In order to simulate evolution steps, individuals are subject to random variations, necessary to explore the problem's solution space.

The basic intent of EAs is to implement the Darwinian concept of survival of the fittest, applying it to functions to optimize. Through each generation, the solutions considered weak in terms of the specific criteria adopted for the optimization face extinction, whereas the best ones combine to produce new individuals that potentially can improve the convergence to an optimal solution.

The first attempts to mimic evolution by simulating genetic processes date back to Fraser [30] and Bremermann [31]. The main contributor, however, is John Holland, who in 1975 published "Adaptation in Natural and Artificial Systems [32]" in which he introduced the main fundamental concepts of genetic algorithms (GA). In GA, each individual of the population has two characteristics: a chromosome and a fitness value representing its quality. The chromosome is composed of genes; in the original formulation each gene was considered as a bit, therefore either 1 or 0, and the chromosome was a string of zeros and ones. In the following years, several researchers developed new forms of GAs.

A chromosome can be viewed as a sorted string or vector. The evolutionary process starts when all fitness values of the initial population have been assigned. Afterwards, the selection process begins, in which some individuals are selected in order to be included in the mating pool. The fittest individuals are more likely to be selected and spread their properties to the *offsprings*: individuals in the mating pool are combined to produce new hybrids whose fitnesses are evaluated to decide whether or not to pass onto the next population, replacing other individuals. It is common practice to keep a constant number of individuals inside a population at each stage.

GA have not been applied practically for system reliability problems until the 1990s, when researchers such as Coit & Smith [33, 34] implemented it in a combinatorial reliability

design problem. The evolutionary optimization proved very efficient in terms of cost-effectiveness of the selection of the parts and allocation redundancies for system reliability. Several authors then tackled optimization problems by GA. For example, Painton and Campbell [35] presented a model based on such methods, highlighting again their robustness and capability of finding the optimum over a high dimensional nonlinear space in a considerably shorter time than the required one for enumeration. In order to improve the reliability of a personal computer, they identified the main components and their failure modes in order to determine some possible improvement levels.

With regards to applications, reliability allocation to minimize total operating costs, subject to an overall plant safety goal, was presented by Yang et al. [36]. System optimization was used to enhance the design, operation and safety of new and/or existing nuclear power plants. They determined the reliability characteristics of reactor systems, subsystems, major components and plant procedures in accordance with a set of top-level performance goals. The cost for improving and/or degrading the reliability of the system was also included in the reliability allocation process as a multi-objective problem formulation. GA was demonstrated to determine effective solutions for a typical pressurized water reactor.

Konak et al. [37] presented general guidelines regarding the implementation of GA for multi-objective reliability optimization, proposing a list of techniques and highlighting the advantages and difficulties of each of them. The reliable network design problem has been studied using multi-objective GA. Kumar et al. [38] presented a multi-objective GA to optimize telecommunication networks while simultaneously improving network performance and design costs given a system reliability constraint. Kim and Gen [39] studied bicriteria spanning tree networks considering the objectives of cost and reliability, while Marseguerra et al. [40] determined optimal surveillance test intervals using a multi-objective GA to improve reliability and availability.

Problems studied by Martorell et al. [41, 42] involved the selection of technical specifications and maintenance activities at nuclear power plants to increase reliability, availability and maintainability for safety-related equipment. They also considered the optimal allocation of more reliable equipment, testing and maintenance activities to assure high reliability, availability and maintainability levels for safety-related systems. Additional limited

resources (e.g., budget and workforce) were required, to form another a multi-objective problem. Solutions were obtained by using both single-objective GA and multi-objective GA, to solve the problem of testing and maintenance optimization with the objective functions of unavailability and cost.

Various other meta-heuristics have been used for reliability optimization. For example, Ant Colony Optimization (ACO) is a population-based, general search technique for the solution of difficult combinatorial problems [43]. The method is inspired by the pheromone trail laying behavior of real ant colonies. In ACO, artificial ants probabilistically build solutions by taking into account pheromone trails, which change dynamically at run-time, to reflect the agents acquired search experience and heuristic information on the problem instance. ACO algorithms have been applied for the reliability optimization of series-parallel systems [44], also including quantity discounts on the redundant components [45], and network optimization by embedding a Cellular Automata approach combined with Monte Carlo simulation for network availability assessment [46], within a multi-objective ACO search engine [47]. ACO has also been applied in hybrid form with Simulated Annealing (SA), called ACO SA, for the design of communication networks [48], where the design problem is to find the optimal network topology for which the total cost is a minimum and the all-terminal reliability is not less than a given level of reliability.

SA is another type of meta-heuristics introduced by Kirkpatrick et al [49] and Cerny [50] as a general probabilistic method for solving combinatorial optimization problems. SA searches the global optimal solution avoiding entrapment in poor local optima by allowing a (probabilistically) occasional uphill move to worse solutions. A SA algorithm for communication network reliability optimization has been proposed [51], which selects the optimal set of links that maximizes the overall reliability of the network subject to a cost constraint, given the allowable node-link incidences, the link costs and the link reliabilities. The algorithm employs a variation of the SA approach coupled with a hierarchical strategy to achieve the global optimum. SA has also been applied to search the optimal solution of system reliability-redundancy allocation problems [52] also considering nonlinear resource constraints [53]. Different SA strategies have been applied to solve multi-objective system reliability optimization problems [54].

Particle Swarm Optimization (PSO) is another algorithm conceptually based on the social

behavior of biological organisms that move in groups, such as birds and fishes [55]. The basic element of PSO is a particle, which can fly throughout the search space toward an optimum by using its own information and that provided by other particles within its neighborhood. As in GA, the performance of a particle is determined by its fitness that is assessed by calculating the objective functions of the problem to be solved. Then, PSO has certainly some similarities to evolutionary algorithms such as GAs, but it also incorporates a cooperative approach. Indeed, all individuals (particles) which are allowed to survive change their positions over time and one particle's successful adaptation is shared and reflected in the performance of its neighbors. Originally developed for the optimization of continuous unconstrained functions, PSO did not attract much attention from the reliability community, because most reliability optimization problems are of discrete nature and have constraints. However, it has been shown that properly adapted PSO can be an effective tool for solving some discrete constrained reliability optimization problems [56]. PSO has, then, been applied to solve reliability optimization and RAP of complex systems [57].

Several optimization meta-heuristics have been designed for various optimization applications in reliability engineering, with varying degrees of success. As no meta-heuristic is so versatile to always outperform the other meta-heuristics in all kinds of reliability optimization problems, developing new, good optimization approaches can be very helpful in some specific applications and benefit practitioners providing more options. Overall, some preferences in practice is given to the use of GAs, as they are able to solve both integer reliability problems and mixed-integer reliability problems. Furthermore, their applicability is not limited to series-parallel systems. In many reliability optimization problems, the optimal solutions found by GAs have turned out to be superior to those of the other meta-heuristic methods for both integer reliability problems (in which component reliabilities are given and redundancy allocation is to be decided) and mixed integer reliability problems (in which both the component reliabilities and redundancy allocation are to be decided simultaneously). Therefore, GAs are very competitive and attractive meta-heuristic methods, especially appropriate for design of nonstandard series-parallel systems. In addition, the multiple solutions found by the GA sometimes vary significantly in the component reliabilities and/or redundancy allocation for systems. This offers the design engineer a variety of options from which to choose with only small differences in the system reliability.

4. Era of Pragmatism

After exhausting much of the inventory of reliability optimization problems that could be solved to optimality by mathematically rigorous methods, researchers entered into an era of pragmatism. The driver for this was the need to expand the types of problems to treat, considering more complex systems and more realistic reliability behaviors of the components, without necessarily being able to mathematically prove the optimality of the solutions (although this remains highly desirable).

Original problem formulations that were solved to optimality often adhered to some common assumptions, although not always, including (*i*) active redundancy, (*ii*) perfect switching of redundant components, (*iii*) limitations on mixing functionally equivalent components within a parallel structure, (*iv*) binary behavior of components and systems, and others. These assumptions simplified the problems and optimal solutions could be found, but artificially constraining the problem spaces far from real conditions. Therefore, the usefulness of these methods was limited, and there was a need to analyze systems with more realistic behaviors, including multi-state systems, uncertain systems, realistic forms of redundancy, etc.

For the more realistic and complex problems, the assumptions or model forms required for mathematical programming algorithms could generally not be satisfied. At the same time, more advanced computers and computer processing provided for exhaustive forms of heuristic search. GA and other forms of meta-heuristics were now used predominantly.

4.1 Multi-state Systems

For components and systems used in practice, often a binary state description (functioning or failed) may not be a proper representation of the reliability behavior, because the component and system reliability performance has a range of different levels (Barlow & Wu [58], Hudson & Kapur [59]). However, evaluation of multi-state system (MSS) reliability is more difficult, and potentially mathematically cumbersome.

Levitin et al [60], and Levitin & Lisnianski [61] presented pioneering and influential research models to optimize system design for multi-state systems. Levitin et al [60] determined an optimal system structure, and Levitin & Lisnianski [61] optimized imperfect preventive maintenance intervals. They used a universal generating function (UGF) approach to evaluate

multi-state reliability and a GA to search the solution space to determine the best solution, although not guaranteed to be the optimal solution. UGF is a convenient function based on a *z*-transform that is useful to systematically and efficiently compute multi-state reliability.

Levitin and his colleagues continued to extend their innovative work to analyze additional applications of multi-state systems. The first formulation of joint structure and maintenance optimization problem for multi-state systems was presented in Levitin & Lisnianski [62], and the optimization approach was extended to systems with common cause failures by Levitin [63]. Later, Levitin & Xing [64] analyzed systems with propagating failures. Each of these research efforts represented fundamental advancements. System reliability optimization could, then, be applied to an entirely new class of systems design problems.

In recent years multi-state models for system reliability assessment have become increasingly popular. In particular, significant research efforts have been devoted to the solution of RAPs for series-parallel multi-state systems (MSSPS) [3, 65-67], which was first introduced in [68]. Series-parallel structures are typically considered because they are quite common in practice. Due to the difficulty of the problem, meta-heuristics are often used to solve MSSPS RAP, even though they can become time-consuming, especially on large systems.

On the other hand, theoretical analysis of meta-heuristics for MSSPS RAP has been generally lacking. Exact/approximated algorithms or guidance for meta-heuristics design have not yet been proposed in the MSSPS RAP literature, while it is important because the application of RAP to multi-state models often requires exhaustive computational resourses. Indeed, the difficulty of solving MSSPS RAP is not only due to the well-known problems of MSS reliability evaluation, but also to the discrete, probabilistic and nonlinear nature of RAP problems.

Another form of a system where the components exhibit multiple states is when component failure time distributions or state probabilities and replaced with a stochastic degradation process. This problem can be particularly challenging when the individual component degradation processes are dependent or have interactions. Song et al [69] determined optimal replacement intervals and inspection intervals for systems with dependent failure processes. Bian and Gebraeel [70] developed a prognostics model for a multi-component system with degradation interactions.

4.2 Uncertainty

The optimization of system reliability relies on a model to provide a representation of the system failure behavior. The model is built on a number of hypotheses on the types of distributions which the stochastic failure processes of the components obey. The values of the parameters of these distributions need to be estimated, and there is always some level of estimation uncertainty. There is intrinsic uncertainty and incomplete knowledge of the system behavior. Uncertainty can be model or structural uncertainty, which exists on the hypotheses of the model, or parameter uncertainty, which exists on the values of its parameters.

In the literature, a number of aspects, factors and causes of uncertainty have been identified, as summarized by Armacost and Pet-Edwards [71], Zimmermann [72]:

- *Lack of information or knowledge*: Lack of information, knowledge and/or data is the main source of uncertainty. This type of uncertainty can be reduced by collecting more information and data.

- *Approximation*: Any model involves some degree of approximation, which is necessary when there is insufficient information to describe exhaustively the phenomenon of interest or when it is desirable to simplify the analysis due to computational constraints or other reasons.

- *Abundance of information or knowledge*: People are incapable of assimilating many pieces of data and information simultaneously. The analyst usually focuses on those parameters and those pieces of data and information that are considered to be more important, while neglecting the others. This type of uncertainty is related to biases in subjective probability assignments (see Kahneman and Tversky [73] and Aven [74]).

- *Conflicting nature of pieces of information/data*: When there is conflicting data, increasing the amount of available information and data would not decrease the uncertainty. More data may just increase the conflict among different pieces of information and data. Some information are affected by errors creating the conflict, although the analyst cannot identify them, or otherwise, the model used by the analyst is poor.

- *Measurement errors*: The measurement of a physical quantity, such as temperature, weight, length, is always affected by the precision of the measurement capability.

- *Linguistic ambiguity*: An expert may express that something is big, but the meaning of "big"

is ambiguous, and can be interpreted in different ways.

- *Subjectivity of analyst judgments*: There can be different interpretations of the same information and data, depending on cultural background and competence of the analyst.

Uncertainty analysis involves identifying and studying the sources of uncertainty and propagating the effects onto the output of the model. Uncertainty analysis determines the uncertainty in the model output that results from uncertainty in the model inputs (Helton et al. [75]). In the practice of reliability engineering and quantitative risk analysis, it is common to distinguish between aleatory and epistemic uncertainty (Apostolakis [76], Helton and Oberkampf [77]). Aleatory uncertainty refers to phenomena occurring randomly, so probabilistic modeling is appropriate to describe such occurrences. Epistemic uncertainty involves quantifying the degree of belief of the analysts on how well it represents the actual system. It is typically expressed as subjective probabilities of the parameters of the probability models. It can be reduced by gathering information and data to improve the knowledge on the system behavior.

For system reliability optimization, uncertainty must be properly accounted for. It is often important to consider the uncertainty in the system reliability estimation so that risky solutions with unsatisfactorily high reliability estimation uncertainty can be avoided. System designers and users are generally risk-averse. Decision makers would generally prefer the design of a system whose reliability is estimated with large confidence, as assured by the low uncertainty of its estimation. Thus, maximization of the system reliability and minimization of its estimation uncertainty is an important formulation, that should be emphasized.

System reliability optimization research originally did not consider the uncertainty in the reliability estimation, although Rubinstein et. al. [78] is an early example of a model to maximize the expectation of a series-parallel system reliability estimate with component uncertainty. However, maximization of the expectation of the reliability estimate may not be adequate, if it is important to avoid system designs with unacceptably high uncertainty. It is therefore desirable to use a multiple-objective optimization algorithm, which explicitly considers the component uncertainty.

In Marseguerra et al. [79], a multi-objective GA is developed to select optimal network designs that balance the dual objectives of high system reliability and low uncertainty in its estimation. Monte Carlo simulation is used to evaluate the objective function and Pareto

optimality is introduced to handle the multiple preference criteria. The decision variables are the numbers of components, x_{ij}, of a given type j to be allocated in the various sections (node pairs & links) i of a network system, $i = 1, 2,…, s$, and $j = 1, 2, …, m_i$. The network is designed to maximize the expectation of the network reliability and minimize the variance of the estimate (by maximizing the negative variance). Introducing cost and weight constraints, the multi-objective optimization problem may be formulated as follows:

Problem P

$$\max\{ \mathrm{E}[\hat{R}(\mathbf{x})], -Var(\hat{R}(\mathbf{x}))\}$$

subject to:

$$\sum_{i=1}^{s}\sum_{j=1}^{m_i} c_{ij} x_{ij} \leq C \;,\; \sum_{i=1}^{s}\sum_{j=1}^{m_i} w_{ij} x_{ij} \leq W$$

This is an appropriate formulation for a risk-averse decision maker, as opposed to most optimization algorithms that require or assume risk-neutrality. Many decision makers may prefer a risk-averse solution, with a marginally lower expected value of reliability compared to a solution with a higher expected value, but with unacceptable uncertainty. Epistemic uncertainty has also been accounted for using interval and fuzzy multi-state models [67, 80-82].

4.3 Different Types of Redundancy

The original formulations of the system reliability optimization problems assumed that all redundancy was active redundancy. This is a convenient assumption because the failure time of a parallel subsystem of components is the maximum of individual component failure times, and the reliability, or probability of survival for some mission time, can be expressed using standard probability principles that are independent of any failure time distribution assumptions. However, many actual subsystem design problems, use a variety of active, cold, warm or hot standby, often within the same design, and therefore the original formulations and solution methods were not practical or applicable for many actual problems.

System designs with active redundancy have fully activated components that can continue to provide needed design functions in the advent of failure of a primary component, until all redundant components have failed as well. Cold standby redundancy involves the use of non-activated components that can be switched-on in response to failure of a primary component. Often systems are designed with both types of redundancy within different parts of

the system, and there are examples where the redundancy type is also a design variable. Cold standby redundancy requires switches to detect failure and activate redundant units. However, the switches can also fail and must be considered in the optimization model. It is assumed that components in cold-standby do not fail, while components in warm standby can fail, but at a lower rate those comparable active or hot standby components. Components in hot standby still require a switching mechanism, but fail at the same rate as active components.

A solution methodology was developed to determine optimal design configurations for nonrepairable series-parallel systems with cold-standby redundancy by Coit [83], who considered a components with non-constant component hazard functions and imperfect switching. There were multiple component choices available for each subsystem and component failure times are distributed according to an Erlang distribution. Optimal solutions are determined using IP with 0-1 decision variables.

There are other engineering system design projects the choice of redundancy type becomes an additional design variable. System design optimization was demonstrated in Coit [84] to maximize reliability when either active or cold-standby redundancy can be selectively chosen for individual subsystems. Formulation of the problem allowing a choice of redundancy strategies is more realistic and practical. Optimal solutions to the problem are found using IP considering imperfect switching of standby redundant components [84]. The optimal system design is distinctly different from the corresponding design obtained with only active or only cold standby redundancy. The same problem was later solved using a GA [85]. Most recent research on systems with imperfect switches has been done by Kim [8] and Levitin et al [86].

The problem with a choice of redundancy strategies has been extended in several original ways including a mixed strategy [87] combining both active and cold-standby components within the same subsystem. Other recent meaningful system reliability optimization research considering mixed component redundancy have provided important models for more varied and practical applications [88-90]. There have also been multiple objective formulations to the problem with different redundancy types [91, 92].

More recently, other different redundancy strategies or types of problems have been considering including the standby element sequencing problem and the combined sequencing and inspection/checkpointing/allocation policy optimization for different types of standby (hot,

cold, warm, mixed). Some research efforts combining system structure optimization with optimal standby component sequencing are included in [8, 9, 93, 94].

5. Era of Active Reliability Improvement

There are currently some very promising on-going research activities that can be considered collectively as an era of active reliability improvement. System reliability optimization is not a static model, but it is being conducted continuously in response to new data being collected on failures, component and system degradation, environmental stresses and usage histories. As an integral part of the optimization process, system performance can be optimized and improved dynamically as new data is collected and analyzed to provide a better understanding of usage conditions and failure behavior, or to compensate for changing conditions. Modern sensor and communications technologies facilitate the collection and transmission of data so that the optimal system design and maintenance plans can be continually enhanced.

Standard assumptions for most system reliability optimization models have been that component failure times form a homogeneous population, and that the failure time distributions, or reliability for a fixed time, are static or stationary. In practice, both assumptions are at best approximations of actual conditions. Although a population of components may form a homogeneous population, their corresponding failure times are influenced by specific environmental stresses and user requirements/stresses that can vary appreciably for specific sub-populations of users or applications. Also, there can be systemic trends in stresses, that can result in shifting failure time distributions over time. In these cases, there is not actually a homogeneous population of identically distributed failure times, and therefore, most optimization models cannot accommodate these realities

The era of mathematical programming resulted in entirely new insights on optimizing system design, and demonstrated how advanced mathematics can be used to solve this problem. The era of pragmatism extended the more theoretical models or developed new ones to address actual conditions of fielded systems. However, in both of these eras, the optimization results were a final result. The solution to these difficult optimization problems was intended to be performed once, perhaps with associated sensitivity analyses. Of course, as new data was collected, the analysis could be repeated, but the optimization process did not directly integrate

changing conditions. In the era of active reliability improvement, the changing conditions and data analyses are an integral part of the model.

5.1 Dynamic System Reliability Models Responding to New Data

Dynamic optimization of system reliability has the potential to achieve responsive system designs, which are highly reliable with changing or diverse conditions. To achieve the highest level of reliability and minimum cost, engineering designs and maintenance plans must address changing conditions or new data that provides better estimates of model coefficients or parameters. To accomplish this, the optimization must be dynamic. The model is solved over-and-over or continually as part of the optimization in response to new data/conditions/etc. as the system is operated.

Yildirim et al [95, 96] present two comprehensive models that involve integrating sensor-based degradation modeling and remaining life distributions with classical mathematical programming, specifically mixed integer programs. The resulting model optimizes predictive maintenance decisions for a complex system to minimize cost. The application being solved is the unit-commitment problem, a well-studied optimization problem, pertaining to power generation and transmission. Yildirim et al. [97] considers opportunistic maintenance scheduling again within an integrated framework that combines mixed integer programming and sensor-based degradation models.

Hao et al [98] addresses dynamic optimization of workload assignment to actively control the degradation and failure time for multiple-units system. Components are degrading and failing, but the rate of degradation is a function of workload assignment. Multiple units are arranged in parallel, and several identical machines may need to operate together to simultaneously produce products to meet the high production demand. This parallel configuration is designed with redundancy to compensate for unexpected events. As data is collected, there is Bayesian updating of degradation model parameters and re-optimization of workloads.

Recent developments by Li et al [99] have proposed providing industrial assets with a degree of agency, in order to enable real time prognostics and optimization of the asset's operation conditions. They consider the feasibility of improving system-level performance in industrial systems by integrating social networks into the IoT (Internet of Things) concept.

Bascifti et al [100] considers a complex system of buses, generators and transmission lines. The model considers the scenarios where unexpected failures happen based on the updated remaining life distributions. The modeling framework in this case is a stochastic optimization model with chance constraints that leverages sensor-based remaining life predictions.

5.2 System Reliability Optimization Customized for Specific Subsets of Users

Data analytics can also be exploited such that the optimal system design can reflect differences within a population. There can be regional differences or fundamental differences within the user population, and by observing and quantifying specific usage conditions and failure patterns, an optimal design can simultaneously correspond to a collection of diverse users or conditions. A failure time distribution can be considered as a function of usage and environmental stresses, and specific reliability values can then naturally vary to reflect these differences.

Ramanan et al [101] studies an advanced distributed optimization problem. There are several interesting aspects and challenges to this problem. The first relates to the computational challenges associated with large scale decentralized optimization and the second relates to the underlying high-performance computing architecture that is would be suitable for such decentralized systems. Advanced data processing of large data sets within subnetworks (local utility companies) is required.

Bei et al. [102] presents a model to fully investigate the integrated redundancy allocation and maintenance planning problem with the presence of uncertain future usage stresses. Component failure time distributions are expressed as a function of environmental and usage stresses. A component system design, with component choices and redundancy levels is selected by the optimization model, but specific preventive maintenance intervals are selected for different usage and environmental stress vectors. The problem is formulated as a two-stage stochastic optimization problem with discrete scenarios defined for different usage and environmental conditions. Zhu et al [103] extends this model by considering uncertain aperiodic changing stresses.

6. Future Challenges in System Reliability Optimization

The safe and productive performance of industrial systems depends on optimal designs

that use equipment reliably, and on testing and maintenance activities that assure the required high level of reliability, availability and maintainability of the equipment. This is done through the efficient assignment of resources that are usually limited. A number of challenges arise in relation to the modern complex systems reliability optimization:

- Integration and response to continual streams of data proving new and updated information.
- Accounting for both aleatory and epistemic uncertainties within the decision-making framework of system reliability optimization
- Cooperative optimization of multi-agent systems, with individual objectives to be optimized within an overall system optimization
- Integrated optimization of reliability design, maintenance, spare parts inventory and logistics management
- Dynamic optimization of evolving systems under changing conditions

7. Conclusions

In this paper, we have provided an organized discussion and review on the evolution of the subject of complex system reliability optimization, which is at the heart of reliability engineering. We have presented how the development of solutions to such problem, and their application, have evolved as a continuum of ideas and techniques, which we have chronologically organized into the three eras of Mathematical Programming, Pragmatism, and Active Reliability Improvement.

In this flow of development and advancement, we have highlighted the joint pull force coming from the fields of operations research and optimization theory, and from the evolution of technology. Their combination has led to the advancement of research ideas and theories, brought new perspectives from the engineering world, and resulted in the development and continuous improvement of methods and techniques to address reliability optimization problems of increasingly complex systems.

The underlying message that emerges from this is that system reliability optimization is an ongoing topic of scientific development and will always be so. The research is actively *pulled* by the advancements in mathematics, operations research and optimization theory, and in response, researchers will continually develop new methods and algorithms to solve, more efficiently than before, broader classes of problems, and larger and more difficult problems. At

the same time, the research is *pushed* by the changes in technology, and in the engineering and social worlds, practitioners will continually demand for new developments to cope with the practical challenges encountered on the field.

In conclusion, today we are treating problems that involve more complex systems and more realistic reliability behaviors of the components, including multi-state, uncertain behaviors, etc. We are beginning to address them dynamically, as new data is collected and analyzed to provide a better understanding of usage conditions and failure behavior, or to compensate for changing conditions, so that the optimal system design and maintenance plans can be continually enhanced. This is possible due to the collection and transmission of data by modern sensor and communications technologies. Yet, new opportunities and challenges are always arising, and it will always be necessary to find efficient ways to solve new problems or problems previously unsolvable.

References

[1]   W. Kuo and V. R. Prasad, "An annotated overview of system-reliability optimization," *IEEE Transactions on Reliability,* vol. 49, no. 2, pp. 176-187, 2000.
[2]   W. Kuo, *Optimal reliability design: fundamentals and applications*. Cambridge University Press, 2001.
[3]   W. Kuo and R. Wan, "Recent advances in optimal reliability allocation," in *Computational intelligence in reliability engineering*: Springer, 2007, pp. 1-36.
[4]   M. S. Chern, "On the Computational-Complexity of Reliability Redundancy Allocation in a Series System," *Operations Research Letters,* vol. 11, no. 5, pp. 309-315, Jun 1992.
[5]   C. Derman, G. J. Lieberman, and S. M. Ross, "On optimal assembly of systems," *Naval Research Logistics Quarterly,* vol. 19, no. 4, pp. 569-574, 1972.
[6]   C. Derman, G. J. Lieberman, and S. M. Ross, "Assembly of systems having maximum reliability," *Naval Research Logistics Quarterly,* vol. 21, no. 1, pp. 1-12, 1974.
[7]   E. El-Neweihi, F. Proschan, and J. Sethuraman, "Optimal allocation of components in parallel–series and series–parallel systems," *Journal of Applied Probability,* vol. 23, no. 3, pp. 770-777, 1986.
[8]   H. Kim, "Maximization of system reliability with the consideration of component sequencing," *Reliability Engineering & System Safety,* vol. 170, pp. 64-72, 2018.
[9]   G. Levitin, L. Xing, and Y. Dai, "Optimization of component allocation/distribution and sequencing in warm standby series-parallel systems," *IEEE Transactions on Reliability,* vol. 66, no. 4, pp. 980-988, 2017.
[10]   R. E. Bellman, "The Theory of Dynamic Programming," The RAND Corporation, 1954.
[11]   R. E. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton University Press, 1957.
[12]   R. E. Bellman and E. Dreyfuss, "Dynamic Programming and Reliability of Multicomponent Devices," *Operations Research,* vol. 6, pp. 200-206, March-April, 1958.
[13]   R. E. Bellman and E. Dreyfuss, *Applied Dynamic Programming*. Princeton, NJ: Princeton University Press, 1962.

[14]   D. E. Fyffe, W. W. Hines, and N. K. Lee, "System Reliability Allocation and a Computational Algorithm," *IEEE Transactions on Reliability,* vol. R-17, pp. 64-69, June, 1968.

[15]   Y. Nakagawa and S. Miyazaki, "Surrogate Constraints Algorithm for Reliability Optimization Problems with 2 Constraints," *IEEE Transactions on Reliability,* vol. 30, no. 2, pp. 175-180, 1981.

[16]   G. B. Dantzig, A. Orden, and P. Wolfe, "The generalized Simplex Method for Minimizing a Linear Form under Linear Inequality Restraints," *The Pacific Journal of Mathematics,* vol. 5, no. 2, 1955.

[17]   P. M. Ghare and R. E. Taylor, "Optimal Redundancy for Reliability in a Series System," *Operations Research,* vol. 17, pp. 838-847, September, 1969.

[18]   R. L. Bulfin and C. Y. Liu, "Optimal Allocation of Redundant Components for Large Systems," *IEEE Transactions on Reliability,* vol. 34, no. 3, pp. 241-247, 1985.

[19]   D. W. Coit and J. C. Liu, "System reliability optimization with k-out-of-n subsystems," *International Journal of Reliability, Quality and Safety Engineering,* vol. 7, no. 02, pp. 129-142, 2000.

[20]   K. B. Misra and U. Sharma, "An Efficient Algorithm to Solve Integer-Programming Problems Arising in System-Reliability Design," *IEEE Transactions on Reliability,* vol. 40, no. 1, pp. 81-91, Apr 1991.

[21]   M. Gen, K. Ida, Y. Tsujimura, and C. E. Kim, "Large-Scale 0-1 Fuzzy Goal Programming and its Application to Reliability Optimization Problem," *Computers & Industrial Engineering,* vol. 24, no. 4, pp. 539-549, Oct 1993.

[22]   M. Gen, K. Ida, and J. U. Lee, "A computational algorithm for solving 0- 1 goal programming with GUB structures and its application for optimization problems in system reliability," *Electronics and Communications in Japan (Part III: Fundamental Electronic Science),* vol. 73, no. 12, pp. 88-96, 1990.

[23]   R. B. Corotis and A. M. Nafday, "Application of mathematical programming to system reliability," *Structural Safety,* vol. 7, no. 2-4, pp. 149-154, 1990.

[24]   J. Song and A. Der Kiureghian, "Bounds on system reliability by linear programming," *Journal of Engineering Mechanics,* vol. 129, no. 6, pp. 627-636, 2003.

[25]   Y. Chang and Y. Mori, "A Study of System Reliability Analysis Using Linear Programming," *Journal of Asian Architecture and Building Engineering,* vol. 13, no. 1, pp. 179-186, Jan 2014.

[26]   G. Levitin, *The Universal Generating Function in Reliability Analysis and Optimization*. Springer, 2005.

[27]   C. A. Floudas, *Nonlinear and mixed-integer optimization: fundamentals and applications*. Oxford University Press, 1995.

[28]   F. A. Tillman, C.-L. Hwang, and W. Kuo, "Optimization Techniques for System Reliability with Redundancy - A Review," *IEEE Transactions on Reliability,* vol. 26, no. 3, pp. 148-155, 1977.

[29]   F. A. Tillman, C.-L. Hwang, and W. Kuo, "Determining component reliability and redundancy for optimum system reliability," *IEEE Transactions on Reliability,* vol. 26, no. 3, pp. 162-165, 1977.

[30]   A. S. Fraser, "Simulation of genetic systems by automatic digital computers I. Introduction," *Australian Journal of Biological Sciences,* vol. 10, no. 4, pp. 484-491, 1957.

[31] H. J. Bremermann, "Optimization through evolution and recombination," *Self-organizing Systems,* vol. 93, p. 106, 1962.

[32] J. H. Holland, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press, 1992.

[33] D. W. Coit and A. E. Smith, "Penalty guided genetic search for reliability design optimization," *Computers & Industrial Engineering,* vol. 30, no. 4, pp. 895-904, Sep 1996.

[34] D. W. Coit and A. E. Smith, "Reliability optimization of series-parallel systems using a genetic algorithm," *IEEE Transactions on Reliability,* vol. 45, no. 2, pp. 254-266, Jun 1996.

[35] L. Painton and J. Campbell, "Genetic Algorithms in Optimization of System Reliability," *IEEE Transactions on Reliability,* vol. 44, no. 2, pp. 172-178, Jun 1995.

[36] J.-E. Yang, M.-J. Hwang, T.-Y. Sung, and Y. Jin, "Application of genetic algorithm for reliability allocation in nuclear power plants," *Reliability Engineering & System Safety,* vol. 65, no. 3, pp. 229-238, 1999.

[37] A. Konak, D. W. Coit, and A. E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," *Reliability Engineering & System Safety,* vol. 91, no. 9, pp. 992-1007, 2006.

[38] R. Kumar, P. P. Parida, and M. Gupta, "Topological design of communication networks using multiobjective genetic optimization," in *Evolutionary Computation, 2002. CEC'02. Proceedings of the 2002 Congress on*, 2002, vol. 1, pp. 425-430: IEEE.

[39] J.-R. Kim and M. Gen, "A genetic algorithm for solving Bicriteria network topology design problems," *Journal of Japan Society for Fuzzy Theory and Systems,* vol. 12, no. 1, pp. 43-54, 2000.

[40] M. Marseguerra, E. Zio, and L. Podofillini, "Optimal reliability/availability of uncertain systems via multi-objective genetic algorithms," *IEEE Transactions on Reliability,* vol. 53, no. 3, pp. 424-434, 2004.

[41] S. Martorell, A. Sánchez, S. a. Carlos, and V. Serradell, "Alternatives and challenges in optimizing industrial safety using genetic algorithms," *Reliability Engineering & System Safety,* vol. 86, no. 1, pp. 25-38, 2004.

[42] S. Martorell *et al.*, "RAMS+ C informed decision-making with application to multi-objective optimization of technical specifications and maintenance using genetic algorithms," *Reliability Engineering & System Safety,* vol. 87, no. 1, pp. 65-75, 2005.

[43] M. Dorigo, G. D. Caro, and L. M. Gambardella, "Ant algorithms for discrete optimization," *Artificial Life,* vol. 5, no. 2, pp. 137-172, 1999.

[44] M. Nourelfath and N. Nahas, "Ant colonies in optimization of system reliability," in *Probabilistic Safety Assessment and Management*, 2004, pp. 1636-1641: Springer.

[45] P. R. McMullen, "Ant-Colony Optimization for the System Reliability Problem with Quantity Discounts," *American Journal of Operations Research,* vol. 7, no. 02, p. 99, 2017.

[46] E. Zio, L. Podofillini, and V. Zille, "A combination of Monte Carlo simulation and cellular automata for computing the availability of complex network systems," *Reliability Engineering & System Safety,* vol. 91, no. 2, pp. 181-190, 2006.

[47] E. Zio and V. Zille, "Multi-objective optimization of network systems by using ANT algorithms," *Quality Technology & Quantitative Management,* vol. 4, no. 2, pp. 211-224, 2007.

[48] B. Dengiz, F. Altiparmak, and O. Belgin, "Design of reliable communication networks: A hybrid ant colony optimization algorithm," *IIE Transactions,* vol. 42, no. 4, pp. 273-287, 2010.

[49] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science,* vol. 220, no. 4598, pp. 671-680, 1983.

[50] V. Černý, "Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm," *Journal of Optimization Theory and Applications,* vol. 45, no. 1, pp. 41-51, 1985.

[51] M. M. Atiqullah and S. S. Rao, "Reliability optimization of communication networks using simulated annealing," *Microelectronics & Reliability,* vol. 33, no. 9, pp. 1303-1319, 1993.

[52] B. D. Mori, H. F. de Castro, and K. L. Cavalca, "Application of Simulated Annealing Algorithm in the Reliability Optimization of Redundant System," SAE Technical Paper 0148-7191, 2004.

[53] H.-G. Kim, C.-O. Bae, and D.-J. Park, "Reliability-redundancy optimization using simulated annealing algorithms," *Journal of Quality in Maintenance Engineering,* vol. 12, no. 4, pp. 354-363, 2006.

[54] B. Suman, "Simulated annealing-based multiobjective algorithms and their application for system reliability," *Engineering Optimization,* vol. 35, no. 4, pp. 391-416, 2003.

[55] J. Kennedy, "Particle swarm optimization," in *Encyclopedia of Machine Learning*: Springer, 2011, pp. 760-766.

[56] G. Levitin, X. Hu, and Y.-S. Dai, "Particle swarm optimization in reliability engineering," in *Computational Intelligence in Reliability Engineering*: Springer, 2007, pp. 83-112.

[57] S. Pant, D. Anand, A. Kishor, and S. B. Singh, "A particle swarm algorithm for optimization of complex system reliability," *International Journal of Performability Engineering,* vol. 11, no. 1, pp. 33-42, 2015.

[58] R. E. Barlow and A. S. Wu, "Coherent systems with multi-state components," *Mathematics of Operations Research,* vol. 3, no. 4, pp. 275-281, 1978.

[59] J. C. Hudson and K. C. Kapur, "Reliability analysis for multistate systems with multistate components," *AIIE Transactions,* vol. 15, no. 2, pp. 127-135, 1983.

[60] G. Levitin, A. Lisnianski, H. Ben-Haim, and D. Elmakis, "Redundancy optimization for series-parallel multi-state systems," *IEEE Transactions on Reliability,* vol. 47, no. 2, pp. 165-172, 1998.

[61] G. Levitin and A. Lisnianski, "Optimization of imperfect preventive maintenance for multi-state systems," *Reliability Engineering & System Safety,* vol. 67, no. 2, pp. 193-203, 2000.

[62] G. Levitin and A. Lisnianski, "Joint redundancy and maintenance optimization for multistate series–parallel systems," *Reliability Engineering & System Safety,* vol. 64, no. 1, pp. 33-42, 1999.

[63] G. Levitin, "Incorporating common-cause failures into nonrepairable multistate series-parallel system analysis," *IEEE Transactions on Reliability,* vol. 50, no. 4, pp. 380-388, 2001.

[64] G. Levitin and L. Xing, "Reliability and performance of multi-state systems with propagated failures having selective effect," *Reliability Engineering & System Safety,* vol. 95, no. 6, pp. 655-661, 2010.

[65] R. Gupta and M. Agarwal, "Penalty guided genetic search for redundancy optimization in multi-state series-parallel power system," *Journal of Combinatorial Optimization,* vol. 12, no. 3, p. 257, 2006.

[66] G. Levitin and A. Lisnianski, *Multi-state System Reliability. Series on Quality, Reliability and Engineering Statistics, vol. 6.* New York: World Scientific Publishing, 2003.

[67] C.-Y. Li, X. Chen, X.-S. Yi, and J.-Y. Tao, "Interval-valued reliability analysis of multi-state systems," *IEEE Transactions on Reliability,* vol. 60, no. 1, pp. 323-330, 2011.

[68] G. Levitin, A. Lisnianski, and D. Elmakis, "Structure optimization of power system with different redundant elements," *Electric Power Systems Research,* vol. 43, no. 1, pp. 19-27, 1997.

[69] S. Song, D. W. Coit, Q. Feng, and H. Peng, "Reliability Analysis for Multi-Component Systems Subject to Multiple Dependent Competing Failure Processes," *IEEE Transactions on Reliability,* vol. 63, no. 1, pp. 331-345, 2014.

[70] L. Bian and N. Gebraeel, "Stochastic modeling and real-time prognostics for multi-component systems with degradation rate interactions," *IIE Transactions,* vol. 46, no. 5, pp. 470-482, 2014.

[71] R. L. Armacost and J. Pet-Edwards, "Integrative risk and uncertainty analysis for complex public sector operational systems," *Socio-Economic Planning Sciences,* vol. 33, no. 2, pp. 105-130, 1999.

[72] H.-J. Zimmermann, "An application-oriented view of modeling uncertainty," *European Journal of Operational Research,* vol. 122, no. 2, pp. 190-198, 2000.

[73] A. Tversky and D. Kahneman, "Judgment under uncertainty: Heuristics and biases," *Science,* vol. 185, no. 4157, pp. 1124-1131, 1974.

[74] T. Aven, "Some reflections on uncertainty analysis and management," *Reliability Engineering & System Safety,* vol. 95, no. 3, pp. 195-201, 2010.

[75] J. C. Helton, J. D. Johnson, C. J. Sallaberry, and C. B. Storlie, "Survey of sampling-based methods for uncertainty and sensitivity analysis," *Reliability Engineering & System Safety,* vol. 91, no. 10-11, pp. 1175-1209, 2006.

[76] G. E. Apostolakis, "How useful is quantitative risk assessment?," *Risk Analysis,* vol. 24, no. 3, pp. 515-520, 2004.

[77] J. C. Helton, J. D. Johnson, and W. L. Oberkampf, "An exploration of alternative approaches to the representation of uncertainty in model predictions," *Reliability Engineering & System Safety,* vol. 85, no. 1-3, pp. 39-71, 2004.

[78] R. Y. Rubinstein, G. Levitin, A. Lisnianski, and H. Ben-Haim, "Redundancy optimization of static series-parallel reliability models under uncertainty," *IEEE Transactions on Reliability,* vol. 46, no. 4, pp. 503-511, 1997.

[79] M. Marseguerra, E. Zio, L. Podofillini, and D. W. Coit, "Optimal Design of Reliable Network Systems in Presence of Uncertainty," *IEEE Transactions on Reliability,* vol. 54, no. 2, pp. 243-253, 2005.

[80] S. Destercke and M. Sallak, "An extension of universal generating function in multi-state systems considering epistemic uncertainties," *IEEE Transactions on Reliability,* vol. 62, no. 2, pp. 504-514, 2013.

[81] Y. Ding and A. Lisnianski, "Fuzzy universal generating functions for multi-state system reliability assessment," *Fuzzy Sets and Systems,* vol. 159, no. 3, pp. 307-324, 2008.

[82] Y.-F. Li, Y. Ding, and E. Zio, "Random fuzzy extension of the universal generating function approach for the reliability assessment of multi-state systems under aleatory and

epistemic uncertainties," *IEEE Transactions on Reliability,* vol. 63, no. 1, pp. 13-25, 2014.

[83] D. W. Coit, "Cold-standby redundancy optimization for nonrepairable systems," *IIE Transactions,* vol. 33, no. 6, pp. 471-478, Jun 2001.

[84] D. W. Coit, "Maximization of system reliability with a choice of redundancy strategies," *IIE Transactions,* vol. 35, no. 6, pp. 535-543, Jun 2003.

[85] R. Tavakkoli-Moghaddam, J. Safari, and F. Sassani, "Reliability optimization of series-parallel systems with a choice of redundancy strategies using a genetic algorithm," *Reliability Engineering & System Safety,* vol. 93, no. 4, pp. 550-556, 2008.

[86] G. Levitin, L. Xing, and Y. Dai, "Mission cost and reliability of 1-out-of-*N* warm standby systems with imperfect switching mechanisms," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 44, no. 9, pp. 1262-1271, 2014.

[87] M. A. Ardakan and A. Z. Hamadani, "Reliability optimization of series–parallel systems with mixed redundancy strategy in subsystems," *Reliability Engineering & System Safety,* vol. 130, pp. 132-139, 2014.

[88] H. Gholinezhad and A. Z. Hamadani, "A new model for the redundancy allocation problem with component mixing and mixed redundancy strategy," *Reliability Engineering & System Safety,* vol. 164, pp. 66-73, 2017.

[89] G. Levitin, L. Xing, and Y. Dai, "Cold vs. hot standby mission operation cost minimization for 1-out-of-*N* systems," *European Journal of Operational Research,* vol. 234, no. 1, pp. 155-162, 2014.

[90] M. Sadeghi and E. Roghanian, "Reliability optimization for non-repairable series-parallel systems with a choice of redundancy strategies: Erlang time-to-failure distribution," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability,* vol. 231, no. 5, pp. 587-604, 2017.

[91] A. Chambari, S. H. A. Rahmati, and A. A. Najafi, "A bi-objective model to optimize reliability and cost of system with a choice of redundancy strategies," *Computers & Industrial Engineering,* vol. 63, no. 1, pp. 109-119, 2012.

[92] J. Safari, "Multi-objective reliability optimization of series-parallel systems with a choice of redundancy strategies," *Reliability Engineering & System Safety,* vol. 108, pp. 10-20, 2012.

[93] G. Levitin, L. Xing, and Y. Dai, "Optimizing Dynamic Performance of Multistate Systems With Heterogeneous 1-Out-of-N Warm Standby Components," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* 2016.

[94] M. Valaei and J. Behnamian, "Allocation and sequencing in 1-out-of-*N* heterogeneous cold-standby systems: Multi-objective harmony search with dynamic parameters tuning," *Reliability Engineering & System Safety,* vol. 157, pp. 78-86, 2017.

[95] M. Yildirim, X. A. Sun, and N. Z. Gebraeel, "Sensor-driven condition-based generator maintenance scheduling—part II: incorporating operations," *IEEE Transactions on Power Systems,* vol. 31, no. 6, pp. 4263-4271, 2016.

[96] M. Yildirim, X. A. Sun, and N. Z. Gebraeel, "Sensor-driven condition-based generator maintenance scheduling—Part I: Maintenance problem," *IEEE Transactions on Power Systems,* vol. 31, no. 6, pp. 4253-4262, 2016.

[97] M. Yildirim, N. Z. Gebraeel, and X. A. Sun, "Integrated Predictive Analytics and Optimization for Opportunistic Maintenance and Operations in Wind Farms," *IEEE Transactions on Power Systems,* vol. 32, no. 6, pp. 4319-4328, 2017.

[98]   L. Hao, K. Liu, N. Gebraeel, and J. Shi, "Controlling the Residual Life Distribution of Parallel Unit Systems Through Workload Adjustment," *IEEE Transactions on Automation Science and Engineering,* vol. 14, no. 2, pp. 1042-1052, 2017.

[99]   H. Li, A. K. Parlikad, and A. Salvador Palau, "A Social Network of Collaborating Industrial Assets," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability,* 2018.

[100]   B. Basciftci, S. Ahmed, N. Gebraeel, and M. Yildirim, "Stochastic Optimization of Maintenance and Operations Schedules under Unexpected Failures," *IEEE Transactions on Power Systems,* no. (in print), 2018.

[101]   P. Ramanan, M. Yildirim, E. Chow, and N. Gebraeel, "Asynchronous Decentralized Framework for Unit Commitment in Power Systems," *Procedia Computer Science,* vol. 108, pp. 665-674, 2017.

[102]   X. Bei, N. Chatwattanasiri, D. W. Coit, and X. Zhu, "Combined Redundancy Allocation and Maintenance Planning Using a Two-Stage Stochastic Programming Model for Multiple Component Systems," *IEEE Transactions on Reliability,* vol. 66, no. 3, pp. 950-962, 2017.

[103]   X. Zhu, X. Bei, N. Chatwattanasiri, and D. W. Coit, "Optimal System Design and Sequential Preventive Maintenance Under Uncertain Aperiodic-Changing Stresses," *IEEE Transactions on Reliability,* 2018.