A Simulation-Optimization Algorithm for Generating Sets of Alternatives Using Population-Based Metaheuristic Procedures

Julian Scott Yeomans

(OMIS Area, Schulich School of Business, York University, Canada)

ABSTRACT: When solving complex stochastic engineering problems, it can prove preferable to create numerous quantifiably good alternatives that provide multiple, disparate perspectives. These alternatives need to satisfy the required system performance criteria and yet be maximally different from each other in the decision space. The approach for creating maximally different sets of solutions is referred to as modelling-to-generate-alternatives (MGA). Simulation-optimization approaches are frequently employed to solve computationally difficult problems containing significant stochastic uncertainties. This paper outlines an MGA algorithm that can generate sets of maximally different alternatives for any simulation-optimization method that employs a population-based procedure. This algorithmic approach is both computationally efficient and simultaneously produces the prescribed number of maximally different solution alternatives in a single computational run of the procedure.

KEYWORDS: Modelling-to-generate-alternatives, Simulation-Optimization, Metaheuristics, Population-based algorithms

Received 28 Aug., 2019; Accepted 13 Sept., 2019 © The author(s) 2019. Published with open access at www.questjournals.org

I. INTRODUCTION

Stochastic engineering decision-making situations typically involve complex design components that can be difficult to incorporate into mathematical programming formulations and are often overwhelmed with numerous unquantifiable specifications [1], [2], [3], [4], [5]. While “optimal” solutions can be calculated for the modelled formulations, they generally do not provide the best solution to the “real” problem as there are usually unmodelled components not apparent when the mathematical models are constructed [1], [2], [6]. Generally, it is more desirable to create a small number of distinct alternatives that permit complementary viewpoints for the problem [3], [7]. These options should be near-optimal with respect to the specified objective(s), but should be maximally different from each other within the decision region. The approach for creating maximally different sets of solutions is referred to as modelling-to-generate-alternatives (MGA) [6], [7], [8].

MGA techniques necessitate a systematic examination of the solution space in order to produce a set of alternatives that are considered good when measured within the modelled objective space but as different as possible from each other in the decision space. The resultant set of solutions should provide alternative perspectives that perform similarly with respect to the modelled objectives, yet very differently with respect to potentially unmodelled features [5]. Decision-makers must conduct a subsequent comparison of the alternatives to determine which alternative(s) most nearly satisfies their specific requirements. Consequently, unlike the more straightforward approach of explicit solution determination inherent in most “hard” optimization methods, MGA approaches are necessarily classified into the decision support realm.

Early MGA algorithms employed direct, incremental approaches for constructing their alternatives by iteratively re-running their procedures whenever new solutions needed to be generated [6], [7], [8], [9], [10]. These iterative approaches replicated the seminal MGA technique of Bril et al. [8] where, once the initial mathematical formulation has been optimized, all supplementary alternatives are produced one-at-a-time. Therefore, these approaches all employed n+1 iterations of their respective algorithms – firstly to optimize the original problem, then to construct each of the n subsequent alternatives [7], [11], [12], [13].

In this paper, it is demonstrated how a set of maximally different solution alternatives can be generated by extending several earlier MGA techniques to stochastic optimization ([12], [13], [14], [15], [16], [17], [18]). In this study, a new stochastic algorithm provides an MGA process that can be accomplished by any population-based mechanism. This new algorithm advances earlier concurrent procedures ([13], [15], [16], [17], [18]) by...
permitting the simultaneous generation of \( n \) distinct alternatives in a single computational run. Specifically, to generate \( n \) maximally different alternatives, the algorithm runs exactly the same number of times that a function optimization procedure needs to run (i.e. once) irrespective of the value of \( n \) [19], [20], [21], [22], [23]. Furthermore, a novel data structure is employed that permits simultaneous alternatives to be constructed in a computationally effective way. This data structure facilitates the above-mentioned solution generalization to all population-based methods. Consequently, this stochastic MGA algorithmic approach proves to be extremely computationally efficient.

II. MODELLING TO GENERATE ALTERNATIVES

Mathematical optimization has focused almost entirely on constructing single optimal solutions to single-objective problems or determining sets of noninferior solutions for multi-objective formulations [2], [5], [8]. While such approaches may create solutions to the mathematical models, whether these outputs are the best solutions to the “real” problems remains can be debatable [1], [2], [6], [8]. Within most “real world” decision-making environments, there are countless system requirements and objectives that will never be explicitly apparent or included in the model formulation stage [1], [5]. Furthermore, most subjective aspects remain unavoidably unmodelled and unquantified in the constructed decision models. This regularly occurs where final decisions are constructed based not only on modelled objectives, but also on more subjective stakeholder goals and socio-political-economic preferences [7]. Several incongruent modelling dualities are discussed in [6], [8], [9], and [10].

When unmodelled issues and unquantified objectives exist, non-conventional methods are needed to not only search the decision region for noninferior sets of solutions, but to also explore the decision region for alternatives that are obviously sub-optimal for the problem modelled. Namely, any search for alternatives to problems known or suspected to contain unmodelled components must concentrate not only on a non-inferior set of solutions, but also necessarily on an explicit exploration of the problem’s inferior solution space.

To demonstrate the consequences of an unmodelled objective in a decision search, assume that the quantifiably optimal solution for a single-objective, maximization problem is \( X^* \) with a corresponding objective value \( Z1^* \). Now suppose that a second, unmodelled, maximization objective \( Z2 \) exists that subjectively incorporates some unquantifiable “politically acceptable” component. Now assume that some solution, \( X^* \), belonging to the 2-objective noninferior set, exists that represents a potentially best compromise solution for the decision-maker if both objectives had somehow been simultaneously evaluated. While \( X^* \) could reasonably be considered as the best compromise solution for the real problem, in the quantified mathematical model it would appear inferior to solution \( X^* \), since it must be the case that \( Z1^* \geq Z1^* \). Therefore, when unmodelled components are incorporated into a decision-making process, mathematically inferior options to the modelled problem could actually be optimal for the real underlying problem. Consequently, when unmodelled issues and unquantified objectives potentially exist, alternative solution procedures are needed to not only explore the decision region for noninferior sets of solutions, but also to concurrently search the decision region for inferior solutions to the problem modelled. Population-based algorithms permit concurrent searches throughout a decision space and prove to be particularly proficient solution methods.

The primary task of MGA is to create a workable set of options that are quantifiably good when measured by all objectives, yet as different as possible from each other within the solution domain. This resulting set of options should produce truly different perspectives that perform similarly with respect to the known modelled objective(s) yet very differently with respect to any unmodelled aspects. By creating these good-but-different solutions, the decision-makers can then examine potentially desirable qualities within the options that may be able to address potentially unmodelled objectives to varying degrees of stakeholder tolerability.

To motivate the MGA process, it is necessary to more formally characterize the mathematical definition of its goals [6], [7]. Assume that the optimal solution to an original mathematical model is \( X^* \) with corresponding objective value \( Z^* = F(X^*) \). The resultant difference model can then be solved to produce an alternative solution, \( X \), that is maximally different from \( X^* \):

Maximize \( \Delta(X, X^*) = \min \{ |X_i - X_i^*| \} \)  
Subject to: \( X \in D \)  
\( |F(X) - Z^*| \leq T \)  

where \( \Delta \) represents an appropriate difference function (shown in (1) as an absolute difference) and \( T \) is a tolerance target relative to the original optimal objective value \( Z^* \). \( T \) is a user-specified limit that determines what proportion of the inferior region needs to be explored for acceptable alternatives. This difference function concept can be extended into a difference measure between any set of alternatives by replacing \( X^* \) in the objective of the maximal difference model and calculating the overall minimum absolute difference (or some

Corresponding Author: Julian Scott Yeomans
other function) of the pairwise comparisons between corresponding variables in each pair of alternatives – subject to the condition that each alternative is feasible and falls within the specified tolerance constraint. The population-based MGA procedure to be introduced is designed to generate a pre-determined small number of close-to-optimal, but maximally different alternatives, by adjusting the value of T and solving the corresponding maximal difference problem instance by exploiting the population structure of the metaheuristic. The survival of solutions depends upon how well the solutions perform with respect to the problem’s originally modelled objective(s) and simultaneously by how far away they are from all of the other alternatives generated in the decision space.

III. SIMULATION-OPTIMIZATION FOR STOCHASTIC OPTIMIZATION

Finding optimal solutions to large stochastic problems proves complicated when numerous system uncertainties must be directly incorporated into the solution procedures ([24], [25], [26], [27]). Simulation-Optimization (SO) is a broadly defined family of stochastic solution approaches that combines simulation with an underlying optimization component for optimization ([24]). In SO, all unknown objective functions, constraints, and parameters are replaced by simulation models in which the decision variables provide the settings under which simulation is performed.

The general steps of SO can be summarized in the following fashion ([25], [28]). Suppose the mathematical model of the optimization problem contains n decision variables, \( X \), represented in the vector \( X = [ X_1, X_2, ..., X_n ] \). If the objective function is expressed by \( F \) and the feasible region is designated by \( D \), then the mathematical programming problem is to optimize \( F(X) \) subject to \( X \in D \). When stochastic conditions exist, values for the objective and constraints can be determined by simulation. Any solution comparison between two different solutions \( X_1 \) and \( X_2 \) requires the evaluation of some statistic of \( F \) modelled with \( X_1 \) compared to the same statistic modelled with \( X_2 \) ([24], [29]). These statistics are calculated by simulation, in which each \( X \) provides the decision variable settings employed in the simulation. While simulation provides a means for comparing results, it does not provide the mechanism for determining optimal solutions to problems. Hence, simulation cannot be used independently for stochastic optimization.

Since all measures of system performance in SO are stochastic, every potential solution, \( X \), must be calculated through simulation. Because simulation is computationally intensive, an optimization algorithm is employed to guide the search for solutions through the problem’s feasible domain in as few simulation runs as possible ([26], [29]). As stochastic system problems frequently contain numerous potential solutions, the quality of the final solution could be highly variable unless an extensive search has been performed throughout the entire feasible region. A stochastic SO approach contains two alternating computational phases; (i) an “evolutionary” module directed by some optimization (frequently a metaheuristic) method and (ii) a simulation module ([30]). Because of the stochastic components, all performance measures are necessarily statistics calculated from the responses generated in the simulation module. The quality of each solution is found by having its performance criterion, \( F \), evaluated in the simulation module. After simulating each candidate solution, their respective objective values are returned to the evolutionary module to be utilized in the creation of ensuing candidate solutions. Thus, the evolutionary module aims to advance the system toward improved solutions in subsequent generations and ensures that the solution search does not become trapped in some local optima. After generating new candidate solutions in the evolutionary module, the new solution set is returned to the simulation module for comparative evaluation. This alternating, two-phase search process terminates when an appropriately stable system state (i.e. an optimal solution) has been attained. The optimal solution produced by the procedure is the single best solution found throughout the course of the entire search process ([30]).

Population-based algorithms are conducive to SO searches because the complete set of candidate solutions maintained in their populations permit searches to be undertaken throughout multiple sections of the feasible region, concurrently. For population-based optimization methods, the evolutionary phase evaluates the entire current population of solutions during each generation of the search and evolves from a current population to a subsequent one. A primary characteristic of population-based procedures is that better solutions in a current population possess a greater likelihood for survival and progression into the subsequent population.

It has been shown that SO can be used as a very computationally intensive, stochastic MGA technique ([29], [31]). However, because of the very long computational runs, several approaches to accelerate the search times and solution quality of SO have been examined subsequently [28]. The next section provides an MGA algorithm that incorporates stochastic uncertainty using SO to much more efficiently generate sets of maximally different solution alternatives.

IV. POPULATION-BASED SIMULTANEOUS MGA COMPUTATIONAL ALGORITHM

In this section, a novel data structure is introduced that permits alternatives to be simultaneously constructed in a computationally efficient way that also enables an algorithmic generalization to solution by any
population-based procedure. Suppose that it is desired to be able to produce \( P \) alternatives that each possess \( n \) decision variables and that the population algorithm is to possess \( K \) solutions in total. That is, each solution is to contain one possible set of \( P \) maximally different alternatives. In this representation, let \( Y_k, k = 1, \ldots, K \), represent the \( k\)th solution which is made up of one complete set of \( P \) different alternatives. Namely, if \( X_{kp} \) is the \( p\)th alternative, \( p = 1, \ldots, P \), of solution \( k, k = 1, \ldots, K \), then \( Y_k \) can be represented as

\[
Y_k = [X_{k1}, X_{k2}, \ldots, X_{kp}] .
\]

(4)

If \( X_{kp}, q = 1, \ldots, n \), is the \( q\)th variable in the \( j\)th alternative of solution \( k \), then

\[
X_{kj} = [X_{kj1}, X_{kj2}, \ldots, X_{kJp}] .
\]

(5)

Consequently, an entire population, \( Y \), consisting of \( K \) different sets of \( P \) alternatives can be written in vectorized form as,

\[
Y^* = [Y_0, Y_1, \ldots, Y_K] .
\]

(6)

The following population-based MGA method produces a pre-determined number of close-to-optimal, but maximally different alternatives, by modifying the value of the bound \( T \) in the maximal difference model and using any population-based metaheuristic to solve the corresponding, maximal difference problem. Each solution within the population contains one potential set of \( p \) different alternatives. By exploiting the co-evolutionary solution structure within the metaheuristic, the algorithm collectively evolves each solution toward sets of different local optima within the solution space. In this process, each desired solution alternative undergoes the common search procedure of the metaheuristic. However, the survival of solutions depends both upon how well the solutions perform with respect to the modelled objective(s) and by how far away they are from all of the other alternatives generated in the decision space.

A straightforward process for generating alternatives would be to iteratively solve the maximum difference model by incrementally updating the target \( T \) whenever a new alternative needs to be produced and then re-running the algorithm. This iterative approach would parallel the original Hop, Skip, and Jump (HSJ) MGA algorithm of Brill et al. [8] in which, once an initial problem formulation has been optimized, supplementary alternatives are systematically created one-by-one through an incremental adjustment of the target constraint to force the sequential generation of the suboptimal solutions. While this approach is straightforward, it requires a repeated execution of the optimization algorithm [7], [12], [13].

To improve upon the stepwise alternative approach of the HSJ algorithm, a concurrent MGA technique was subsequently designed based upon the concept of co-evolution ([13], [15], [17]). In a co-evolutionary approach, pre-specified stratified subpopulation ranges within an algorithm’s overall population were established that collectively evolved the search toward the creation of the specified number of maximally different alternatives. Each desired solution alternative is represented by each respective subpopulation and each subpopulation undergoes the common processing operations of the procedure. The survival of solutions in each subpopulation depends simultaneously upon how well the solutions perform with respect to the modelled objective(s) and by how far away they are from all of the other alternatives. Consequently, the evolution of solutions in each subpopulation toward local optima is directly influenced by those solutions contained in all of the other subpopulations, which forces the concurrent co-evolution of each subpopulation towards good but maximally distant regions within the decision space according to the maximal difference model [7]. Co-evolution is also much more efficient than a sequential HSJ-style approach in that it exploits the inherent population-based searches to concurrently generate the entire set of maximally different solutions using only a single population [12], [17].

While a concurrent approach can exploit population-based solution approaches, the co-evolution process can only occur within each of the stratified subpopulations. Consequently, the maximal differences between solutions in different subpopulations can only be based upon aggregate subpopulation measures. Conversely, in the following simultaneous MGA algorithm, each solution in the population contains exactly one entire set of alternatives and the maximal difference is calculated only for that particular solution (i.e. the specific alternative set contained within that solution in the population). Hence, by the evolutionary nature of the population-based search procedure, in the subsequent approach, the maximal difference is simultaneously calculated for the specific set of alternatives considered within each specific solution – and the need for concurrent subpopulation aggregation measures is circumvented.

Using the terminology introduced above, the steps in the stochastic MGA procedure are as follows ([14], [19], [20], [21], [22], [23], [32]). It should be apparent that the stratification approach outlined in this algorithm can be easily modified to accommodate any population-based solution procedure.

Preliminary Step. In this initialization step, solve the original optimization problem to determine the optimal solution, \( X^* \). As with prior solution approaches ([13], [15], [16], [17], [18]) and without loss of generality, it is entirely possible to forego this step and construct the algorithm to find \( X^* \) as part of its solution processing. However, such a requirement increases the number of computational iterations of the overall procedure and the initial stages of the processing focus upon finding \( X^* \) while the other elements of each population solution remain essentially “computational overhead”. Based upon the objective value \( F(X^*) \),

Corresponding Author: Julian Scott Yeomans
establish P target values. P represents the desired number of maximally different alternatives to be generated within prescribed target deviations from the $X^*$. Note: The value for P has to have been set a priori by the decision-maker.

Step 1. Create the initial population of size K in which each solution is divided into P equally-sized partitions. The size of each partition corresponds to the number of variables for the original optimization problem. $X_{kp}$ represents the $p^{th}$ alternative, $p = 1,...,P$, in solution $Y_k$, $k = 1,...,K$.

Step 2. In each of the K solutions, evaluate each $X_{kp}$, $p = 1,...,P$, using the simulation module with respect to the modelled objective. Alternatives meeting their target constraint and all other problem constraints are designated as feasible, while all other alternatives are designated as infeasible. A solution can only be designated as feasible if all of the alternatives contained within it are feasible.

Step 3. Apply an appropriate elitism operator to each solution to rank order the best individuals in the population. The best solution is the feasible solution containing the most distant set of alternatives in the decision space (the distance measure is defined in Step 5). Note: Because the best solution to date is always retained in the population throughout each iteration, at least one solution will always be feasible. A feasible solution for the first step can always consist of P repetitions of $X^*$.

Step 4. Stop the algorithm if the termination criteria (such as maximum number of iterations or some measure of solution convergence) are met. Otherwise, proceed to Step 5.

Step 5. For each solution $Y_k$, $k = 1,..., K$, calculate $D_k$, a distance measure between all of the alternatives contained within the solution. As an illustrative example for determining a distance measure, calculate

$$D_k = \Delta (X_{ka}, X_{kb}) = \text{Min} \mid X_{ka} - X_{kb} \mid, \quad a = 1,...,P, \quad b = 1,...,P, \quad q = 1,...,n.$$  

This represents minimum absolute distance between all of the alternatives contained within solution k. Alternatively, the distance measure could be calculated by some other appropriately defined function.

Step 6. Rank the solutions according to the distance measure $D_k$ objective – appropriately adjusted to incorporate any constraint violation penalties for infeasible solutions. The goal of maximal difference is to force alternatives to be as far apart as possible in the decision space from the alternatives of each of the partitions within each solution. This step orders the specific solutions by those solutions which contain the set of alternatives which are most distant from each other.

Step 7. Apply appropriate metaheuristic “change operations” to each of the solutions within the population and return to Step 2.

V. CONCLUSIONS

“Real world” decision-making situations inherently involve complicated performance components that are further confounded by incongruent requirements and unquantifiable performance objectives. These decision environments frequently contain incompatible design specifications that are problematic – if not impossible – to incorporate when ancillary decision support models are constructed. Invariably, there are unmodelled elements, not apparent during model formulation, that can significantly affect the adequacy of its solutions. These confounding features require the decision-makers to integrate numerous uncertainties into their solution process before an ultimate solution can be determined. Faced with such inconsistencies, it is unlikely that any single solution can simultaneously satisfy all ambiguous system requirements without significant compromises. Therefore, any decision support approach must somehow address these complicating facets in some way, while simultaneously being flexible enough to condense the potential effects within the intrinsic planning incongruities.

This paper has provided an updated stochastic MGA algorithm that directs stochastic SO search processes. This new computationally efficient approach establishes how population-based algorithms can simultaneously construct entire sets of close-to-optimal, maximally different alternatives by exploiting the evolutionary characteristics of any population-based solution method. This MGA approach simultaneously creates several solutions containing the requisite problem features, with each alternative generated providing a very different perspective to the problem considered. The practicality of this stochastic MGA approach can be readily extended into numerous disparate applications and can be clearly modified to suit many “real world” planning situations. Such extensions will be explored in future research.

REFERENCES

[1]. Brugnach, M., A. Tagg, F. Keil and W.J. De Lange, Uncertainty matters: computer models at the science-policy interface. Water Resources Management, 2007, 21: p. 1075-1090.
[2]. Janssen, J.A.E.B., M.S. Krol, R.M.J. Schielen and A.Y. Hoekstra, The effect of modelling quantified expert knowledge and uncertainty information on model-based decision making. Environmental Science and Policy, 2010, 13(3): p. 229-238.
[3]. Matthies, M., C. Giupponi and B. Ostendorf, Environmental decision support systems: Current issues, methods and tools. Environmental Modelling and Software, 2007, 22(2): p. 123-127.
Simulation-Optimization Algorithm for Generating Sets of Alternatives Using Population-Based Metahuristic Procedures

A Simulation-Optimization Algorithm for Generating Sets of Alternatives Using Population-Based Metahuristic Procedures

Julian Scott Yeomans

Corresponding Author: Julian Scott Yeomans

Page 6 | Page 6