ENGINEERING APPLICATIONS OF HEURISTIC MULTILEVEL OPTIMIZATION METHODS

Jean-François M. Barthelemy
NASA Langley Research Center
Hampton, VA
Abstract Some engineering applications of heuristic multilevel optimization methods are presented and the discussion focuses on the dependency matrix that indicates the relationship between problem functions and variables. Decompositions are identified with dependency matrices that are full, block diagonal and block triangular with coupling variables. Coordination of the subproblem optimizations is shown to be typically achieved through the use of exact or approximate sensitivity analysis. Areas for further development are identified.

Introduction

Ever since optimization methods have been applied in engineering, practitioners have attempted to use them in multilevel schemes. These are procedures where a large problem is broken down into a number of smaller subproblems; this phase is referred to as decomposition. These subproblems are optimized separately and an iterative process is then devised which accounts for the coupling so that when it is converged, the resulting optimum is that of the original non-decomposed problem; this phase is referred to as coordination.

Multilevel methods can be classified as formal or heuristic according to whether the decomposition and the coordination phases are exclusively based on the mathematical form of the problem or on understanding of the underlying physics. In general, formal methods are more amenable to convergence studies than heuristic methods. The distinction between the two classes of methods is somewhat arbitrary, however, and, depending on how it is presented, a method may be shown to belong to either class.

This paper covers applications of heuristic multilevel optimization methods in engineering design. Problems are assumed to be formulated as static nonlinear parametric programming problems. While most applications are for structural design problems, reference will be made also to selected papers in mechanical, power and electrical engineering.

The paper begins with a review of the objectives of multilevel optimization and a description of typical applications. The two following sections address the decomposition problem and the coordination problem. The paper concludes with an assessment of the state-of-the-art and recommendations for further work. While the paper discusses primarily two-level formulations, most methods may be adapted to decompositions with more than two levels. For the sake of generality, the presentation remains in terms of a generic design problem. Only a limited number of representative papers will be cited.

Objectives and Examples of Application

Some design problems naturally have a multilevel structure as the calculation of their constraints or objective functions are themselves the results of minimization or maximization problems. Haftka [1] showed that the design of damage tolerant space trusses and wing boxes can be formulated with a constraint on maximum collapse load.

By far, the most commonly cited reason for resorting to multilevel optimization is the improvement of the numerical performance of optimization algorithms. In structural optimization, early attempts were direct extensions of the fully stressed design methodology. Using methods devised
by Giles [2] and Sobieszczanski and Loendorf [3]. Fulton et al. [4] designed a complete aircraft model that involved on the order of 700 design variables and 2500 constraints. Schmit and Mehrinfar [5] followed with optimization of truss and wing box models that included local and global constraints while Hughes [6] developed similar ideas for naval structures. Using a method first proposed by Sobieszczanski [7], Wrenn and Dovi [8] optimized a fairly complex transport wing model with 1200 variables and 2500 nonlinear constraints. Substructuring has also been used to decompose optimization problems. Nguyen [9] used it to reduce the cost of the sensitivity analysis phase. Schmit and Chang [10] and Svensson [11] have looked at optimizing substructures independently. In other engineering applications, multilevel approaches were used to design underground energy storage systems (Sharma, [12]), speed reducers (Datseris, [13]), microwave systems (Bandler and Zhang, [14]) and to solve the optimum power flow problem (Contaxis et al. [15]).

Formulating a multilevel problem can also be used to improve its mathematical conditioning since variables that have different orders of magnitudes and rates of change can be kept separate in the optimization process. Probably the most common example of such application is the simultaneous sizing and optimization of the geometry of structures in which the sizing problem is solved for fixed geometry in an inner loop, while in the outer loop, the geometry is modified to optimize the design. This approach has been used primarily for space trusses and frameworks, examples are given by Felix [16]. Kirsch [17] used a similar formulation to conduct the simultaneous analysis and optimization of reinforced concrete beams.

The design of complex engineering systems is by nature multilevel. Designers carry out the effort by breaking the total problem into subproblems and assigning each to different units of the engineering team. Each unit has developed its own design methodologies and successful designs result from skillful integration of objectives, requirements and constraints from each unit. This becomes a coordination problem. Sobieszczanski [7] was the first to propose to use multilevel coordination methods to solve multidisciplinary design problems. Rogan and Kolb [18] showed how a transport aircraft preliminary design problem can be treated as a multilevel optimization problem.

**Decomposition**

The general form of the original, non-decomposed optimization problem is as follows (vectors are boldfaced and scalars use normal script):

\[
\min f(\mathbf{X}), \text{ st } g(\mathbf{X}) \leq 0, \ h(\mathbf{X}) = 0
\]

The relationship between variables and functions (objective and constraints) can be described symbolically by the dependency matrix (Fig. 1). There is one column in the matrix for each variable (or vector of similar variables) and one row for each function (or vector of similar functions); the objective function is listed first. Entry \(i,j\) indicates qualitatively the relation between function \(j\) and variable \(i\). In our figures an entry \(X\) indicates function \(i\) depends on variable \(j\); no entry indicates function \(i\) does not depend on variable \(j\). Figure 1 corresponds to Prob. 1, a general nonlinear programming problem where all functions are assumed to depend on all variables.

As discussed by Carmichael [19], "...decomposition implies breaking the system into subsystems with interactions and breaking the problem [variables], constraints and [objective] into [variables], constraints and [objectives] associated with the subproblems. Decoupling... may be carried out by the introduction [or identification] of interaction variables such that there results independent optimization problems at the lower level." Typical approaches to decomposition are discussed below.

**Decomposition of the Variable Vector**

Without any special structure (that is with a fully populated dependency matrix), Prob. 1 may always be decomposed by partitioning the variable vector:
It may then be replaced by \( n \) problems, the \( i \)th of which is

\[
\min_{x_i} f(x_i, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n), \quad \text{st } g(x_i, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \leq 0,
\]

and

\[
h(x_i, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = 0
\]

where an overbar on a variable indicates that the variable is held fixed.

This approach has been used for simultaneous configuration optimization and sizing (Lev, [20]) and optimal load flow control (Contaxis et al. [15]). Typically, no real decoupling results from such a decomposition (the dependency matrix remains fully populated), unless one of the subproblems can be further decomposed as in Kirsch [17] or Vanderplaats et al. [21].

### Block-Diagonal Dependency Matrix

From the standpoint of decomposition, a problem having an additively separable objective function and a dependency matrix as in Fig. 2a (assuming suitable re-ordering of the variables and constraints) is ideal, since it yields totally uncoupled subproblems which can be solved independently of each other. The original problem formulation reads:

\[
\min_{x=x_1, \ldots, x_n} f(x) = \sum_{i=1}^{n} f_i(x_i) \quad \text{st } g_i(x_i) \leq 0, \quad h_i(x_i) = 0 \quad i=1, n
\]

resulting in \( n \) independent subproblems:

\[
\min_{x_i} f_i(x_i) \quad \text{st } g_i(x_i) \leq 0, \quad h_i(x_i) = 0
\]

While design problems seldom have such form, it is often assumed that they have a similar form in which some functions depend strongly on some variables and only weakly on others. This situation is described in Fig. 2b where dots denote weak dependency. Assuming additively separable objective function, this yields the following \( n \) subproblems:

\[
\min_{x_i} f_i(x_i, x_{i-1}, x_{i+1}, \ldots, x_n), \quad \text{st } g_i(x_i, x_{i-1}, x_{i+1}, \ldots, x_n) \leq 0,
\]

and

\[
h_i(x_i, x_{i-1}, x_{i+1}, \ldots, x_n) = 0
\]

One of the major shortcomings of this method is that it cannot explicitly handle constraints which strongly depend on variables belonging to different subsystems. Sobieszczanski and Loendorf [3] and Hughes [6] devised an ad hoc procedure to correct the overall design for violations of these constraints.

Generally, the decomposition of the problem is arrived at in a very natural way; it is imposed by the structure or the layout of the engineering system considered. Therefore, very few systematic approaches to decomposition exist. An exception is that used by Datseris [13] for the design of mechanisms. Here the key idea is to divide the set of design variables in mutually exclusive subsets so that some measure of the coupling between the variable subsets is minimized. Coupling is measured by an interdependence function based on the design problem objective function. If a decomposition in two subsets is desired, the first step is to randomly identify two subsets of variables. Then a systematic approach is used to exchange variables among the subsets in an effort to lower the value of the interdependence function.

Another approach to systematic decomposition is proposed by Bandler and Zhang [14] in their optimization of large microwave systems. Their starting point is a matrix similar to the dependency matrix introduced above. They use a matrix whose \( i,j \) entry is the normalized sensitivity derivative of function \( i \) with respect to variable \( j \) (or a sum of sensitivity derivatives calculated at various points in the design space). They manipulate the rows and columns of the matrix to finally identify the subproblem to optimize
starting with the reference function group (with the worst contribution to the objective) and the candidate variable groups (those that affect that reference function group). Optimization proceeds with repeated redefinition of the variable and function groups making up the subproblem, which as the optimum design is reached, includes all variables and functions.

**Block-Angular Dependency Matrix with Coupling Variables**

Reasonably complex engineering design problems cannot typically be formulated with a block-diagonal (Fig. 2a) or even a quasi-block diagonal (Fig. 2b) structure. Indeed, as alluded to before, some constraints depend strongly on variables belonging to several subproblems. A more typical structure is the block-angular structure with coupling variables of Fig. 3a. This may result from the existence of a hierarchical structure in the model in which two levels of variables and functions exist. At the higher level, the higher level (or system or global) variables affect directly the higher level constraints. At the lower level, for fixed higher level variables, the lower level (or subsystem or local) variables affect directly the lower level constraints. Further decoupling may exist that results in a number of independent lower level subproblems. The coupling higher level variables are the interaction variables. Assuming additively separable objective function, the starting problem would be given by

\[
\min_{\mathbf{y}, \mathbf{x}_1, \ldots, \mathbf{x}_n} f_0(\mathbf{y}) + \sum_{i=1}^n f_i(\mathbf{y}, \mathbf{x}_i) \quad \text{st} \quad g_0(\mathbf{y}) \leq 0, \quad g_i(\mathbf{y}, \mathbf{x}_i) \leq 0 \quad i=1, n \\
\text{and} \quad h_0(\mathbf{y}) = 0, \quad h_i(\mathbf{y}, \mathbf{x}_i) = 0 \quad i=1, n
\]  

(7)

The resulting higher level subproblem would then be

\[
\min_{\mathbf{y}} f_0(\mathbf{y}) \quad \text{st} \quad g_0(\mathbf{y}) \leq 0, \quad h_0(\mathbf{y}) = 0
\]  

(8a)

while there would be n independent lower level subproblems:

\[
\min_{\mathbf{x}_i} f_i(\mathbf{y}, \mathbf{x}_i) \quad \text{st} \quad g_i(\mathbf{y}, \mathbf{x}_i) \leq 0, \quad h_i(\mathbf{y}, \mathbf{x}_i) = 0
\]  

(8b)

Haftka [22] gave a penalty formulation for the same initial problem.

To derive a problem structure as in Eq. 7 from a general nonlinear programming problem as described in Eq. 1, equality constraints sometimes need to be introduced. They typically express the consistency between the higher level and the lower level models of the system. These can impede convergence of the process. Thereja [23] proposed to linearize them at each optimization step and to use them to eliminate some variables of the problem and thus reduce its size. Schmit and Merhinfar [5] transformed these equality constraints in penalty-type objective functions for the lower level subproblems allowing for incomplete satisfaction of the equalities at the beginning of the optimization process, and in effect, only achieving a quasi-block-angular structure as in Fig. 3b.

The issue of automatically generating a problem structure as in Eq. 7 for complex engineering systems was first addressed by Rogan and Kolb [18] who suggested handling it as scheduling problem.

**Coordination**

Coordination amounts to devising a scheme iterating among the subproblem optimizations such that the final solution is that of the original problem (or one of its solutions). Central to the coordination process is the identification of coordination variables (Carmichael [19]). These variables are held fixed at the lower level, giving independent subproblems which are solved separately, and then information is returned to the higher level to update the value of the coordination variables. This cycle is repeated until convergence is achieved. Some modification of the higher level subproblem is necessary to ensure coordination.

Applications that rely on variable vector or block-diagonal (or quasi-block-diagonal) decompositions generally do not possess any coordination mechanism. In the former case, coordination is really not necessary since...
each subproblem deals with all the functions of the problem. In the latter case, this lack of coordination has been long known to prevent finding even a local minimum of the problem and probably accounts for some of the disappointing results reported by Svensson [11]. In the context of structural applications, Sobieszczanski [24] indicated: "...Minimization of the individual component masses does not guarantee minimization of the total mass. This situation is caused by the inability to control the load path on the assembled structure level...". Schmit and Chang [10] offer a unique approach to coordinating problems using a substructuring formulation. They write the problem variable vector:

\[ \mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i \]  

Each vector \( \mathbf{x}_i \) is manipulated at the local level to satisfy local constraints while minimizing stiffness (hence boundary force) changes; vector \( \alpha \) is manipulated at the global level to minimize the global objective, satisfy the global constraints and some local constraints that cannot be satisfied at the local level.

Block-angular decompositions with coupling variables provide an explicit coordination mechanism. A feasible coordination technique is always used in which the higher level variables are taken as the coordination variables. Generally, to provide a means of coordination at the higher level, the effect of changes in lower level designs due to changes in higher level variables must be known.

For example, at the end of each lower level optimization, Schmit and Merhinfar [5] update limits on higher level behavioral (dependent) variables to reflect new lower level designs. To coordinate the lower level designs Felix [16] suggests to take a search direction at the higher level that will minimize the system objective function while continuing to satisfy the constraints active at the conclusion of the lower level optimizations. A one dimensional search is performed at the higher level that accounts for possible higher level constraints.

Since lower level optima are obtained for fixed value of the coordination variables, they really are implicit functions of these variables. For the subproblem of Eq. (8b), denoting optimum quantities with an (*), we have

\[ f_1^*(\mathbf{y}, \mathbf{x}_1^*) = f_1^*(\mathbf{y}, \mathbf{x}_1^*(\mathbf{y})) - f_1^*(\mathbf{y}) \]  

Optimization at the higher level must therefore continue in a direction that maintains these lower level optima. To achieve coordination, the problem of Eq. (8a) must then be restated:

\[ \min_{\mathbf{y}} f_0(\mathbf{y}) + \sum_{i=1}^{n} f_i^*(\mathbf{y}) \quad \text{s.t.} \quad g_0(\mathbf{y}) \leq 0, \quad h_0(\mathbf{y}) = 0 \]  

One approach to constructing approximations to the implicit relations of Eq. (10) is to repeat the lower level solutions for several combinations of higher level variables. The resulting information can be used in non-gradient optimization schemes or in gradient schemes with finite-difference-based derivative estimates. Kunar and Chan [25] used the conjugate direction and the conjugate gradient method. In addition to being computationally expensive, this approach is prone to round-off and truncation errors. Alternately, as proposed by Sharma et al. [12] the information can be used in surface-fitting procedures to construct approximate response surfaces giving the lower level optima explicitly as functions of the higher level variables. While this approach appears effective for small problems, the size of the sample necessary for large problems with large numbers of higher level variables will become prohibitive.

Another approach proposed by Sobieszczanski [7], and Sobieszczanski et al. [26] is to resort to sensitivity analysis of optimum solutions. This technique provides exact derivatives of the solution of lower level subproblems with respect to higher level variables and permits the generation of first-order approximations:
Haftka [22] used a similar approach for penalty function formulations.

Complete sensitivity analysis of optimum solutions (variables, objective and constraints) is numerically costly since it requires second-order derivatives of these functions. However, as shown by Barthelemy and Sobieszczanski [27], if only the lower level objectives must be known for the coordination mechanism, the additional calculations are limited to the problem first-order derivatives.

Sensitivity derivatives are also discontinuous functions of higher level variables (Barthelemy and Sobieszczanski, [28]). Presumably, lower level subproblem unconstrained formulations based on penalty function formulations (Haftka [22]) or envelope functions (Sobieszczanski [7]) should eliminate that difficulty. However, as shown by Barthelemy and Riley [29] in the case where envelope functions are used, driving the solution of the approximate unconstrained subproblems to that of the original constrained ones often results in rapidly varying (albeit still continuous) gradients, a phenomenon that all but brings back the derivative discontinuity issue. It is likely that the same problem occurs with penalty functions formulation. Haftka [22] proposed to limit the effect of discontinuity by restricting optimization to one step at each level in each cycle. Vanderplaats and Cai [30] proposed an interesting approach to approximate sensitivity analysis that should anticipate constraint switching. No definitive solution exists for this difficulty, but no example was ever shown where the derivative discontinuity precluded convergence of the procedure.

Concluding Remarks

This brief review shows that heuristic multilevel optimization methods have a demonstrated potential in engineering design. The most promising decomposable problem statement considered is block-diagonal with coupling variables. These variables are used at the higher level of the decomposition to provide for decoupling of the lower level subproblems and coordination of their optimization. The lower level subproblems communicate with the higher level subproblem with sensitivity information that may be based on formal sensitivity analysis. Various schemes have been proposed and some have been demonstrated on very large problems.

Very little work focuses on the decomposition process itself that is on the approach to be taken to obtain such a block angular structure. If multilevel optimization is to be applied to truly large engineering systems, then the ideas of Rogan and Kolb [18] on scheduling must be further developed. One direction is to account not only on the existence of coupling as they have done but also on the strength of coupling between variables and functions as was done by Bandler and Zhan [14].

As stated above, efficiency of the algorithm is one of the most cited reasons to resort to multilevel optimization. Yet few of the results in the literature are concerned with more than convergence of the algorithm. Haftka [22] showed that significant savings could result from limiting iteration of the subproblems to as little as one iteration per cycle, while Thareja and Haftka [23] showed how further gains could be made by exploiting the structure of the problem when calculating and storing derivatives. Barthelemy and Riley [29] and Vanderplaats et al. [21] showed good results combining decomposition and approximations. The works of Bandler and Zhan [14], as well as Barthelemy and Riley [29] indicate that it is worthwhile in each cycle to optimize only those subproblems that have the strongest influence on the problem objective.

Multilevel procedures are ideally suited for execution in parallel. Surprisingly, no engineering application of multilevel methods on parallel processors has ever been implemented. Young [31] demonstrated the feasibility of using Sobieszczanski’s [7] approach on a network of engineering workstations.
Finally, as all methods developed for design, multilevel methods must be made to conform better to the design process itself. Most complex engineering systems require more than two levels for modelization. Initial work by Sobieszczanski et al. [32] and Kirsch [17] should be pursued. Likewise, particularly in the multidisciplinary context, problems are likely to have several objectives. Multilevel/multiobjective formulations are necessary to determine what design is obtained when each discipline-subproblem deals with its own variables, objective and constraints.

References


