One of the more difficult optimal design tasks occurs when the data describing the system to be optimized is either highly nonlinear or noisy or both. This situation arises when trying to design restraint systems for automotive crashworthiness using the traditional lumped parameter analysis methods. The nonlinearities in the response can come from either abrupt changes in the occupants interaction with the interior or from relatively minor fluctuation in the response due to the interactions of two restraint systems such as belts and airbags. In addition the calculated response measures are usually highly nonlinear functions of the accelerations. Two approaches using an approximate problem formulation strategy are proposed. One approach uses a first-order approximation based on finite difference derivatives with a nonlocal step size. The second and more effective approach uses a second-order curve fitting strategy. Successful example problems of up to 16 design variables are demonstrated.

A conservative design strategy using a derivative-based constraint padding is also discussed. The approach proves effective because analytical expressions are available for the second-order terms. © 1995 John Wiley & Sons, Inc.

INTRODUCTION

Although formal optimization methods have been applied to a wide range of problems in mechanical design, very little of this work has been directed toward problems of vehicle crashworthiness. This is partially because this area is somewhat limited in scope, being primarily of interest to the automotive industry. A more pervasive reason, however, is the inherent nonlinearity in this behavior. Not only are there large displacements both in structural and occupant behavior, but these deformations also extend the materials involved, both structural and human, into nonlinear regimes. Contact occurs both in the deforming structure and in the interior as the occupant encounters various objects. Attempts to model this behavior using mathematical modeling methods date from the late 1960s and early 1970s. Traditionally, separate technical communities have been involved with structural modeling and occupant modeling. Two methods are commonly employed by each community. One approach is to use an essentially lumped parameter approach. For structures this approach is typified by Ni and Fine (1978) and is shown in Fig. 1. The major masses are represented by lumped masses and the deformable structure is modeled by nonlinear springs, typically represented as a force deflection table. These curves can either be obtained by test or subsystem simulation. The occupant representations are typified by those described in Prasad and Chou (1989) and shown in Fig. 2. The occupant is represented by a collection of rigid ellipses connected by springs. The interactions of the occupant with objects inside the passenger compartment are handled by mutual force deflection functions obtained from a subsystem test or
FIGURE 1 Lumped mass structural model.

FIGURE 2 Lumped parameter occupant model with air cushion.

FIGURE 3 Vehicle crash severity index (VCSI) response as a function of design variables (Bennett et al., 1977).

Bennett et al. (1977) used a lumped parameter model and a feasible directions algorithm to find feasible designs. Song (1986) and Lust (1992) have used a hybrid approach with a simple approximation strategy to create explicit subproblems that were then handled by a feasible directions algorithm. All of this work has used finite difference methods for calculating the derivatives. These articles also discussed one of the fundamental difficulties with attempting formal optimization approaches with this class of problems. Figure 3, reproduced from Bennett et al. (1977) representing response quantities as functions of a design variable, clearly shows the nonlinearities in the system. Some of the nonlinearities are abrupt enough that the notion of a derivative is questionable. It is not clear that locating the design at a minimum of one of these functions is desirable, because the design may be very sensitive to small changes in the design variable. The crashworthiness area then presents an opportunity to examine and develop optimization strategies for highly nonlinear problems where the response calculations are sufficiently costly that the number of function evaluations needs to be limited. The present work will attempt to address some of these problems. The particular analysis code used is an improved version of the occupant simulation program, CAL3D (Deng, 1990; Wang and Ngo, 1990), that can handle the current passive restraint systems of belts and air cushions and can also be extended to include lumped mass models of the structure. Although the focus of the article is on developing design strategies that are effective for finding robust solutions to highly nonlinear and potentially discontinuous problems, to
make the example problems understandable it will be necessary to briefly discuss the formulation of design problems in crashworthiness.

**DESIGN VARIABLES AND RESPONSE QUANTITIES**

In general, any quantity in the lumped parameter model is available as a design variable, although usually the design variables are selected to be quantities associated with the nonlinear force-deflection curves. These curves could represent the force-deflection behavior of a structural component such as a front rail or the instrument panel that the occupant may encounter. Typically a scaling factor that is applied to the force axis of these curves is chosen as the design variable. This is attractive because usually this change can be physically accomplished in the subsystem design. In addition design variables may be directly related to physical quantities such as the stiffness of a belt, the vent size in an air bag, or the mass flow into the air bag.

Typically crashworthiness response is assessed in terms of several mechanical response measures on the occupant that are regulated by US Government standards in a 30 mph barrier impact. These quantities are the head injury criterion (HIC), which is an integration of the acceleration experienced by the head and has a maximum target of 1000; the maximum chest Gs, which has a maximum target of 60 Gs; and the loads in the femurs, which have a maximum target of 1000-kg force. In previous work (Bennett et al., 1977; Song, 1986; Lust, 1992) only the structure was considered, so intermediate accelerations and displacements were used as response quantities. However, in the present work occupant response quantities are available so they will be used. The selection of an objective function is still not obvious and three different approaches will be illustrated at various times in this report. The first approach suggests that the desire is to decrease the sum of the normalized response quantities that we will call the injury criterion (IC). Constraints would then be placed on the individual criterion.

\[
\text{OBJ} = \text{IC} = \frac{\text{HIC}}{1000} + \frac{\text{chest G}}{60} + \frac{\text{lfemur}}{1000} + \frac{\text{rfemur}}{1000}. \quad (1)
\]

A second alternative is to recognize that each of these quantities does not have equal weight in terms of the harm that the human experiences. Therefore a weighted injury criterion could be created such as

\[
\text{OBJ} = \text{IC} = 0.60 \left( \frac{\text{HIC}}{1000} \right) + 0.35 \left( \frac{\text{Chest G}}{60} \right) + 0.025 \left( \frac{\text{lfemur}}{1000} + \frac{\text{rfemur}}{1000} \right). \quad (2)
\]

The previous discussions have focused on selecting some measure of safety performance as the goal. While this is certainly a desirable goal, the process of engineering usually involves a compromise between the performance of a system and the cost to produce the system. Although we are not yet at the point where a total cost function could be identified for a safety system such as an air cushion system, an argument could be made that some measure of the amount of propellant used in the inflator is a useful indirect measure of the cost efficiency of an air cushion system. Therefore, for an air cushion system an alternative formulation is to minimize the mass flow scaling parameter subject to constraints on the injury measures, including the weighted composite injury criterion given in Eq. (2).

**OPTIMIZATION CONCEPTS**

One of the concepts that has emerged from the work in structural optimization is the idea of creating an approximate model on which the optimization is actually performed. Although not necessary, this concept forms a convenient way to formulate the safety design problem. In this concept the analysis is used to provide information about a current point in the design space, which may include derivative information. Then an explicit, continuous approximate problem is generated, such as first-order Taylor’s series approximation. This approximate problem is then given to a formal optimization routine and an optimum of this approximate problem is obtained. The formal optimization then becomes a straightforward and inexpensive process because any function information required is obtained from the simple, explicit approximations and the form of these approximations is such that the optimization will converge. This new design point is then evaluated using the full analysis method and the cycle is repeated until some completion criteria are reached. In general, each approximate subproblem is allowed to operate in a region of the design space sufficiently close to the current initial point that the approximations retain some validity.
Move limits on the order of 10–50% are typically used. Two implementations of this strategy for the crashworthiness problem will be presented.

**Sequential Linearization**

This approach creates a linearization of the objective function and all constraints for each subproblem that is essentially a first-order Taylor series. Therefore, at each step a function value and the derivatives of each function with respect to each design variable are required. Because currently no analytical derivatives are routinely available for nonlinear time dependent integration programs such as CAL3D, these sensitivities must be obtained by finite difference methods. However, as indicated by Fig. 3, the selection of the finite difference step size may be crucial. If the standard approach of taking a vanishingly small step is used, true local derivatives may be obtained, but the information may give the wrong global direction (Fig. 3). As a result, for this class of problems a moderately large step may be more practical. Another phenomenon that occurs with this method is that during the approximate problem, sufficient movement may take place to step over local minima. Therefore, even if the magnitude of the derivative is inaccurate, as long as the sign is correct the design may proceed correctly.

These benefits lead to the fundamental difficulty with this approach. Unless the problem is highly constrained, the subproblem optimum will lie on the subproblem bounds of the design variables. This, coupled with the relatively large subproblem bounds to permit jumping over local minima, produces a problem formulation that has difficulty in the final convergence process. This is traditionally handled by automatically decreasing the move limits as the design starts to converge. This can lead to a slow and costly convergence process.

**Second-Order Approximations**

The logical extension of the above process would be to create a second-order approximation at each approximate problem step. In general this turns out to be extremely expensive from a computational standpoint because order $n^2$ function (CAL3D) evaluations (where $n$ is the number of design variables) will be required at each step. Even if only the diagonal terms of the second-order matrix are used, order $2n$ evaluations are required at each step. Alternatively, one might consider creating second-order models of the entire design space. This again will require at least $n^2$ function evaluations. A number of approaches have been suggested for implementing higher order approximations, several of which are referenced by Free et al. (1987). The particular approach we have implemented is from Vanderplaats (1979) and is conceptually straightforward.

At any point in the design a second-order approximation is constructed about the current best design by fitting a full second-order approximation using a least squares method to determine the “best” second-order fit for the data. The new optimum determined by the solution of the approximate problem is then added to the list of points and the process is continued until a stopping criterion is reached. The process can be started with any number of points. If at any iteration, less than the number of points required for a full second-order fit are available, a reduced approximation is constructed. Terms are added sequentially through the first-order terms, the diagonal second-order terms, and finally the off diagonal terms. Usually only the last $m$ points are retained so that as the design converges only points close to the optimum are retained. This type of approach would seem to have several advantages for the optimization of occupant response. First, the approximate problem will not retain any of the underlying noise of the analysis as indicated in Fig. 3, but it will retain a second-order description of the objective and constraint functions. Second, if a group of starting points is chosen to span the design space, the design will initiate with a rough global description of the design space. If move limits are chosen to be less than the total design space, the design will tend to move fairly slowly and to fill in additional points that are close to the optimum. Thus, the selection of points will include both local and global information until the maximum number of retained points ($m$) is reached. These approximations will always be quadratic if sufficient second-order information is available, and the success of the process is related to how well the true function can be approximated by a quadratic function over the part of the design space considered. This will most likely preclude falling into narrow steep valleys, or they will be noticeable because the differences between the approximate optimum and the full solution at that point will be significant. As indicated previously, a large number of function values may be required whatever approach is used. The choice of starting values is somewhat
arbitrary; however, several possibilities are obvious. One would be to choose \( n + 1 \) points so that the first-order terms are quite well approximated. The logical choice would be to pick values at the extremes of the ranges for the design variables. Following this line of thinking one could also turn to design of experiments schemes including the Taguchi orthogonal arrays. When considered in this light this approximation is essentially similar to steps that have been proposed for extending experimental design past the first array (Free et al., 1987).

EXAMPLES

The above concepts have been implemented in an optimization capability that uses CAL3D as an analyzer. Either of the above strategies is available as well as any other capability that is available in the ADS program (Vanderplaats et al., 1983). Therefore more traditional approaches such as feasible directions can be tried. The objective function and constraints can be any function of the response quantities that are available through CAL3D. These include the standard measures such as HIC, chest Gs etc., as well as maximum and minimum joint forces, segment forces, and spring displacements.

Example 1

The first example is based on a simulation of an unrestrained occupant and a passenger side inflatable restraint. It is similar to the model shown in Wang and Ngo (1990) and is shown symbolically in Fig. 2. Two design variables were used; the vent area of the air bag from which the gasses escape as the energy is absorbed, which will vary between \( 0.001 \) and \( 20.0 \) cm\(^2\), and a scaling factor on the part of the mass flow into the bag that occurs after the bag is taut and can vary between 0.1 and 3.0. The objective function was that described in Eq. (2) and each individual criterion was constrained to be below its reference value. Three starting conditions were used. In the course of the examples shown, these constraints did not influence the design. This allows a convenient graphic interpretation, because the only significant quantities are the two design variables and the weighted IC. A plot of these quantities and the results of the optimization are shown in Fig. 4. For each value of the vent design variable, the shape of the IC is roughly parabolic. The increased values of the IC for low values of the mass flow rate are due to insufficient gas in the bag that allows contact with the interior of the vehicle. The increased values for high mass flow rates are due to more gas than needed, resulting in a stiffer bag and insufficient energy absorption from the venting of the gas. Also shown on the figure is the second-order fit at the final design. Twenty-seven function evaluations were required. A second run was made with identical starting conditions but with the objective function set to minimize the mass flow parameter (the second design variable) and IC constrained to be below 0.5. The results of this run are also shown in Fig. 4. Twelve function evaluations were required. Clearly for these problems the second-order approach will find the correct solutions, even though there are local nonlinearities in the curves and the design space is somewhat complex.

Example 2

This example combines a front structure model and an air cushion and an unbelted occupant. It is essentially the model in Bennett et al. (1991) and is a combination of the structural model in Fig. 1 and the occupant model in Fig. 2. The structural model has 14 springs. For this example a subset of three of these springs (1, 5, and 8) plus the air bag sensing time and the air bag vent area were chosen as design variables. The full set of 16 design variables is presented as example
3. This subset was determined to be the most important of the full 16 design variables. For the structural springs the design variables were chosen to be the scaling factors on the y or force axis of the component force deflection curves. The initial scaling factor was 1.0 and the range was from 0.5 to 1.5. The vent area could vary from 70 to 85 mm$^2$ and the firing time from 0.015 to 0.035 s. The objective function was a slight modification of Eq. (1)

$$IC = \frac{HIC}{1200} + \frac{CSI}{800} + \frac{(lfemur)}{1000} + \frac{(rfemur)}{1000}$$

and the constraints were placed on each individual criterion to be less than the reference value. CSI is the chest severity index. There is no particular physical reason for selecting the normalizing factors of 1200 on HIC and 800 on CSI. At the initial design HIC was 1659 so that the first constraint was violated; however, for all solutions shown the final design was feasible.

This example was chosen to illustrate how some of the different methods perform. First, results for the linear approximation method and the more traditional feasible directions method (FDM) as embodied in ADS are examined. Both of these methods require gradient information and two different multiplicative factors (0.01 and 0.1) on the design variable to calculate the finite difference step sizes were used. For a finite difference factor size of 0.01, FDM was unable to converge to a lower minimum that was obtained by the other methods because it found a local valley from which it was unable to extricate itself. With a finite difference step factor of 0.1 this problem was not encountered. The linear approximation approach with a finite difference factor of 0.01 indicated a large jump in objective function in the fourth step, which is indicative of an incorrect sign in a derivative, but was able to recover in the fifth step. It is possible that this erratic behavior could continue or the design might continue to converge. On the other hand, when the finite difference factor size was 0.1 the convergence was smoother. Based on this study, a finite difference factor of 0.1 was used, each with all values at the nominal except one placed at its maximum. This is labeled High Starting. Similarly the Low Starting uses the lower bound values for each design variable. The set labeled Taguchi uses a Taguchi array to select the values. Because five variables do not fit exactly into a Taguchi orthogonal array the extra values were chosen to represent some, but not all, of the interactions. The Taguchi High set was chosen only considering the values above the initial design rather than the whole array because the final solution is generally in the lower range of the design variables and the High Starting proved to be a difficult starting point. As can be observed from Fig. 5, with the exception of the High Starting all of the starting groups converged to similar optima. The best options from each approach are shown in Fig. 6. They all converge used, each with all values at the nominal except one placed at its maximum.
to similar solutions; however, the second-order approximation method requires less than half of the number of the function evaluations of the linear approximation and about 25% of those required by FDM.

Example 3

A larger example problem was created by expanding the design variables to include all 14 structural springs giving a total of 16 design variables. The results of these optimizations are given in Fig. 7. The finite difference parameter used in FDM and the linear approximation strategy is 0.1. A set of initial conditions similar to the High Starting was used for the second-order approximations. In this example FDM got stuck in a local minima even though a large finite difference step was used. The other approaches identified similar minima that were not significantly different from that identified in example 2. Again the second-order approximation required approximately half as many function evaluations.

CONSERVATIVE OR ROBUST DESIGNS

The problem formulations shown above produce solutions that are on the boundaries of the design space in some sense. In many problems one or more constraints will be active, such as in the case where the mass flow term is minimized. Alternatively if the design is unconstrained at the optimum the design will be at the absolute minimum of the objective function. There is always a question as to how sensitive the design is to variations in any of the parameters of the problem, including the design variables. The traditional way of handling this uncertainty is through safety factors applied to the performance constraints,

\[ g_i(x) + \text{safety factor} - 1 < 0. \]  

Although this is a simple and straightforward approach, there has always been a desire to use a somewhat more reasoned approach. Bennett and Lust (1990) present a discussion of some of these concepts with respect to structural design problems. The notion of conservativeness or robustness has a particular importance in the safety area. Because of the complex and varied nature in which accidents occur, we find it necessary to identify a subset of these events to characterize the safety of a vehicle. For example, the 30 mph barrier test is used to characterize the frontal impact of a vehicle. When we use analytical techniques in the design process we further simplify these events in the mathematical models, such as CAL3D models, that we employ. Therefore one can argue that it is important in the automated design process to identify designs that are relatively insensitive, or at least with some known sensitivity to the variations possible in the final design. This section will address some approaches to this problem.

For structural optimization problems a sensitivity or gradient based approach to robust design was proposed Bennett and Lust (1990). This approach suggested that the safety factor or padding term be made proportional to the gradient of the constraint.

\[ g_i - 1.0 + \text{pad}_i < 0 \]

\[ \text{pad}_i = \sum \left| \frac{\partial g_i}{\partial x_j} \right| (c_i x_j) \]  

where \( c_i \) becomes the percent variation in the design variable that will be protected against by this representation. The absolute values are necessary to account for worst-case possibilities in the summation. This approach is especially attractive in that we have available an analytical second-order representation of the response functions if we use the second-order approximation capabilities available in the optimization program described previously. There are two possible implementations of this concept available. The first is to use Eq. (5) as it stands and the sensitivity
calculations that will be used in the optimization process then can contain the correct and full second-order information that is available,

$$\frac{\partial}{\partial x_k} \text{pad}_i = c_k \left| \frac{\partial g_i}{\partial x_k} \right| + \sum \text{sgn} \left( \frac{\partial g_i}{\partial x_k} \right) \left( \frac{\partial^2 g_i}{\partial x_k \partial x_k} \right)(c_j x_j) \tag{6}$$

This implementation contains the correct gradient information to identify a direction in the design space that will reduce the first-order sensitivities.

A second implementation would be to retain all of the information including the second-order information, in the new constraint form,

$$\text{pad}_i = \sum \left| \frac{\partial g_i}{\partial x_k} \right| (c_j x_j) + \sum \sum \left| \frac{\partial^2 g_i}{\partial x_k \partial x_k} \right| (c_j x_j) (c_m x_m) \tag{7}$$

The sensitivities now take the form

$$\frac{\partial}{\partial x_k} \text{pad}_i = c_k \left| \frac{\partial g_i}{\partial x_k} \right| + \sum \text{sgn} \left( \frac{\partial g_i}{\partial x_k} \right) \left( \frac{\partial^2 g_i}{\partial x_k \partial x_k} \right)(c_j x_j) + \sum \left| \frac{\partial^2 g_i}{\partial x_k \partial x_k} \right| (c_j c_k x_j) \tag{8}$$

that contains essentially the same directional information as does Eq. (6) because no third-order information is available. Experience has shown that the second formulation [Eqs. (7), (8)] works considerably better. This is because for many of the approximate problems the second-order effect dominates even over the rather small range of the design variables considered in the padding term.

This is potentially a difficult problem to solve. If there are fairly significant second-order effects, the padding terms can become quite large, particularly because of the absolute value effect. Because the functions that are used in the actual optimization are approximations of the true behavior, these padding terms may either overestimate or underestimate the real effect. If they overestimate, they may so constrain the problem that no feasible solution can be found by the optimizer, even though one may exist. On the other hand, if the effect is underestimated the optimizer may find a feasible solution that is found to be infeasible when evaluated by CAL3D.

Results of implementations of this formulation are given in Fig. 8. The example problem is the same as shown in Fig. 4 for minimizing the mass flow with an IC constraint of 0.5. Three cases were run, one with \(c = c_i = 0.05\) (or 5% variation in the design variables), and one with \(c = c_i = 0.1\) shown in Fig. 8, and one with \(c = c_i = 0.15\). For \(c = 0.05\) and 0.10, after each design was obtained four CAL3D runs were made with the four possible combinations of the maximum variations in the design variables. All of these designs proved to be feasible with respect to the original unpadded constraints. These designs are also shown in Fig. 8. The mass flow for the 0.10 padded optimum was 0.77 as opposed to 0.37 for the unpadded design in Fig. 4. Thus, the protection against a 10% variation requires a 0.4 increase in the mass flow parameter objective function. For a 5% variation the increase was 0.19. Design runs with \(c = 0.15\) were also made, but feasible solutions could not be found. That is, there are no designs for which we can guarantee that the design variables could vary by 15% and the resulting design would not exceed the IC constraint. Examination of the graphical representation in Fig. 8 has not identified any design variable combination that would meet these conditions.

**DISCUSSION**

From the example problems shown it appears that the two proposed methods do avoid some of the
local minima that have caused problems for more traditional strategies such as the feasible directions method. In addition fewer function evaluations are required. It also appears that finite difference steps of the order of 0.1 or 10% may prove to be a good starting point for the gradient-based linearization method and that move limits in the 20–30% range are reasonable. Clearly these numbers are problem dependent. Several different sets of starting values seemed to work for the second-order approximation method. Experiments have indicated that starting with approximately \( n + 1 \) values works better than trying to start with a smaller number of values and that if the initial design spans the boundaries of the space, at least some global feel about the first-order sensitivities will be available at the start. For these examples the second-order approximations consistently required approximately half the number of function evaluations that were required by the linearization method. If it is possible to analytically obtain derivatives that give other than local information, this trend would be reversed.

A conservative problem formulation strategy based on the second-order approximation strategy was developed. This proved to be fairly effective in finding designs that would remain feasible under variations in the design variables due to outside influences.

The authors would like to thank R. V. Lust and J. T. Wang for their many helpful suggestions.

REFERENCES


Submit your manuscripts at
http://www.hindawi.com