Taboo Search: An Approach to the Multiple Minima Problem

Djurdje Cvijović and Jacek Klinowski

Described here is a method, based on Glover's taboo search for discrete functions, of solving the multiple minima problem for continuous functions. As demonstrated by model calculations, the algorithm avoids entrapment in local minima and continues the search to give a near-optimal final solution. Unlike other methods of global optimization, this procedure is generally applicable, easy to implement, derivative-free, and conceptually simple.

Many problems in practically all fields of science, technology, technical design, and econometrics involve global optimization—that is, the determination of the global minimum (or maximum) of a function of an arbitrary number of independent variables that may be continuous or discrete. Until relatively recently, global optimization received little attention (*1*, *2*). It appears not to be possible to design methods that offer an absolute guarantee of locating the global minimum for an arbitrary function, and most "stochastic" methods provide an asymptotic guarantee. Simulated annealing (SA), the first such method to become widely known, was introduced in 1983 (*3*, *4*). Despite considerable recent progress, mainly in combinatorial applications (involving functions defined in a discrete domain) (*5*–*7*), the general problem of global optimization remains unsolved. Efficient and reliable methods are thus urgently needed.

In most cases of practical interest, global optimization is very difficult. This is because of the omnipresence of local minima, the number of which tends to increase exponentially with the size of the problem. Global optimization arises in chemistry in connection with the morphology of aggregates of atoms, conformational analysis, and the study of the geometry of interacting molecules (*8*). Conventional minimization techniques, which are time-consuming and tend to converge to whichever local minimum they first encounter, are of limited value in such cases (*9*, *10*). The solution in these cases may not be the global minimum but a local minimum sensitive to the starting point. Also, these methods are unable to continue the search after a local minimum is reached.

The principal requirement before any global optimization method is that it must be able to avoid entrapment in local minima and continue the search to give a near-optimal solution whatever the initial conditions. SA (*3*, *4*, *11*) and the genetic algorithm (GA) (*12*), the most popular techniques in chemical applications, meet this requirement. We suggest here an approach that uses the concept of the taboo search (TS), a stochastic global optimization method originally developed by Glover (*13*, *14*) for very large combinatorial optimization tasks. We extended it to continuous-valued functions and investigated its potential for functions of many variables and many local minima.

Our work has been inspired by the fact that TS is very general and conceptually much simpler than either SA or GA. We believe that this in itself justifies its closer investigation. Further, it has been shown (*13*–*16*) that TS is superior to SA both in the time required to obtain a solution and in the quality of the latter in solving the graph coloring and the traveling salesman problems, the two most challenging combinatorial optimization tasks. In addition, TS is a flexible framework of a variety of strategies originating from artificial intelligence and is therefore open to further improvement.

We define the global optimization problem as follows:

$$\text{minimize } f(s): s \in S$$

where f is the function to be optimized (the "objective function") and S is a set of feasible solutions (the "solution space"), each represented by a point in S. Because maximizing f is equivalent to minimizing $-f$, we limit ourselves to the problem of minimization.

The TS uses a modification of "local search," which starts from some initial solution and attempts to find a better solution. This becomes the new solution, and the process restarts from it. The procedure continues step by step until no improvement is found to the current solution.

The concepts of a "move" and "neighborhood" are common to most heuristic and algorithmic procedures (*11*). In TS, a move is a transition $(s' \rightarrow s'')$ from one trial solution (s') to another (s''). The "move value" is the difference $f(s'') - f(s')$. The move "improves" only if the move value is negative. A neighborhood $N(s)$ of the solution s is the collection of solutions in S to which the algorithm can move from s in a single step.

In an attempt to improve the current solution s^*, TS considers the neighbors of s^*—that is, the elements of $N(s^*)$. $N(s^*)$ is therefore the set of moves that may be applied to s^* at a particular step to produce a new solution. The complete procedure is a sequence of such moves, and the actual definitions of the neighborhood depend on the particular implementation and the nature of the problem.

In order to produce a new solution, the local search algorithm uses only improving moves. SA goes further: besides improving moves, it allows nonimproving moves with a certain acceptance probability under the control of the temperature parameter. Taboo search goes further still: At each iterative step it examines all moves in $N(s^*)$, both improving and nonimproving. In order to avoid a blind search, TS uses a prescribed problem-specific set of constraints, known as "taboo conditions," which must be satisfied for the move to be considered admissible; otherwise, the move is taboo.

A move remains taboo only during the "taboo period," a certain specified number of iterations. The "aspiration condition" is defined to enable certain "interesting" moves. If this is satisfied, a taboo move becomes admissible—that is, the algorithm is allowed to ignore the taboo status of a move temporarily ("strategic forgetting") provided that the move is judged to be interesting. The taboo condition and the aspiration condition together are a heuristic device, a kind of "learning procedure" that benefits from information acquired during previous iterations. Thus, TS performs an intelligent search of the solution space. From among the admissible moves at each iterative step, TS accepts the move with the lowest move value. This move might not lead to a better solution, but enables the algorithm to continue the search without becoming confounded by the absence of improving moves and to "climb out" of local minima.

We introduced a neighborhood structure in continuous space that we call the "conditional neighborhood." The solution space S (hypercube in R^n, where n is a number of variables) is partitioned into disjunct cells by division of the coordinate intervals along the $x_1, x_2, x_3, \ldots x_n$ axes into $p_1, p_2, p_3, \ldots p_n$ parts. The problem-specific empirical partition parameter $P = (p_1, p_2, p_3, \ldots p_n)$ determines a unique partition of S into cells, thus specifying the "address" of each cell. At each iterative step, we draw n_s sample points from a uniform distribution over n_c randomly chosen cells. These points are the neighbors of the current solution s^*. The size of the neighborhood $N(s^*)$ is $n_c n_s$

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

and remains constant while the content of the neighborhood changes at each iteration. At each iterative step, TS accepts a nontaboo move with the lowest move value, not necessarily to a cell that is a neighbor of the current solution.

We used two kinds of taboo moves. (i) At a particular iterative step, a move is considered taboo if it produces a solution that lies in the "taboo region" of the solution space S, consisting of cells visited during L preceding iterations. (ii) The move becomes taboo if it results in a deterioration of the objective function f greater than some specified value.

The management of taboo moves as described by (i) is done by a "taboo list." This circular list is controlled by the parameter L, known as the size of the taboo list, which may be fixed or variable depending on the application or stage of the search. The taboo list is initially empty. At the first, second, and finally the Lth iteration starting from the current solution s', a certain move $s' \rightarrow s''$ from cell $C(s')$ to cell $C(s'')$ is accepted, producing a new solution s''. The address of $C(s'')$ is then added to the taboo list, in sequence from positions 1 to L. At the $(L + 1)$th iteration, the address of the current cell is added at the beginning of the list after its first (that is, oldest) element is removed and the others have been translated.

The management of taboo moves as described by (ii) is done by keeping the track of the worst value of f throughout the computation and maintaining the "elite list" of addresses of the most promising cells. For the initial step, we drew $10n$ sample points

procedure for all proposed optimization methods; the test functions are discussed in greater detail elsewhere (*1*, *2*, *6*). The Hartman family of functions is defined as

$$f(x) = -\sum_{i=1}^{i=4} c_i \exp\left(-\sum_{j=1}^{j=n} a_{ij}(x_j - p_{ij})^2\right)$$

where $0 \leq x_j \leq 1$, with j equal to any value up to n. Two particular tasks, known as H3 and H6, correspond to $n = 3$ and $n = 6$. The values of c_i, a_{ij}, and p_{ij} are given elsewhere (*17*). The Shubert function

$$f(x_1, x_2) = \left\{\sum_{i=1}^{i=5} i \cos[(i + 1)x_1 + i]\right\} \left\{\sum_{i=1}^{i=5} i \cos[(i + 1)x_2 + i]\right\}$$

has 760 local minima in the region $-10 \leq x_1, x_2 \leq 10$, 18 of which are global with $f(x_1, x_2) = -186.7309$. The Rastrigin function has 50 local minima, one of which is global.

Using the test functions listed in Table 1, we compared the performance of TS with the performance of other stochastic global optimization methods (Table 2). Numerical experiments were carried out with the pure random search (PRS), whereas the performance figures for multistart and two different SA algorithms were taken from the literature. The efficiency was quantified in terms of the number of function evaluations necessary to find the global minimum.

Pure random search (*18*) evaluates $f(s)$ at a fixed number of points chosen randomly from a uniform distribution over S. The smallest value of $f(s)$ is then a candidate for the global minimum. Because the aim is to evaluate the efficiency of finding the global minimum, we implemented the PRS algorithm, which terminates once such point has been found with less than 1% error. The average results of four runs are listed in Table 3. The multistart method is a natural extension of PRS. Local search is applied to every point in a sample drawn from a uniform distribution over S, and the local minimum with the lowest function value is considered to be the global minimum.

The TS parameters listed in Table 1 were found during initial tests. The maximum number of iterations allowed without improvement was 100. The results of continuous TS optimization of test functions (Table 3) were the average outcome of 100 independent runs. The reliability was excellent: in each case at least 90% of the runs

Table 1. Standard test functions and TS parameters used for global optimization.

Name of function	Function abbreviation	Dimension (n)	Partition parameter P	(n_c, n_s)	L
Goldstein-Price	GP	2	(6, 6)	(1, 6)	6
Branin	BR	2	(6, 6)	(1, 6)	6

were successful (with the final result within 2 to 3% of the global minimum). With this degree of precision, the global minimum in all our test functions was "isolated" from local minima, so that the solution could always be refined to any desired accuracy by any local optimizer. We used the findminimum routine, a built-in local optimizer in Mathematica (Wolfram Research, Champaign, Illinois), which we believe to be a variant of the Newton-Raphson method.

Together with SA and GA, TS has been singled out as extremely promising for future practical applications (*19*). This work is a first step in applying the ideas and strategies of TS to continuous optimization. SA, GA, and TS are all based on a combination of combinatorial optimization and concepts from rather unlikely fields. Thus, SA is inspired by statistical physics and in essence amounts to a numerical simulation of the physical annealing of solids where, by slowly decreasing the temperature from the molten state, the system solidifies in a state of minimum energy. Genetic algorithm relies on the Darwinian principle of evolution: The algorithm cross-breeds trial solutions and allows only the "fittest" to survive after several iterations. Taboo search stems from the general tenets of intelligent problem-solving and is based on concepts from artificial intelligence.

Preliminary results for a standard set of test functions thus indicate that continuous TS is reliable and efficient, even more so than PRS and the multistart method. Taboo search reduces the amount of blind search that is characteristic of earlier techniques. The results obtained with TS compare favorably with those obtained with SA. The efficient optimization of the Hartman family of functions is particularly encouraging, given the great physical importance of the sums of Gaussian functions.

Our approach to continuous global optimization has several attractive features: (i) TS avoids entrapment in local minima and continues the search to give a near-optimal final solution; (ii) it is problem-independent and can be applied to a wide range of tasks; (iii) it does not require any information about the derivatives of the function to be minimized; (iv) it is very easy to implement and the entire procedure occupies only a few lines of code; and (v) it is conceptually much simpler than SA and GA. For instance, instead of using the metropolis algorithm, choosing a cooling program, and specifying an annealing schedule (the initial and final temperatures, temperature decrement, and the length of the Markov chain), the elementary version of TS (involving one taboo list and an aspiration function) requires only empirical control parameters: the size of the taboo list and a parameter defining the partition of the solution space.

All the same, we are aware that no stochastic method can be guaranteed to solve the multiple minima problem in a finite number of steps, and any method may require long computing times. Also, just like SA, TS is a heuristic method and thus requires theoretical justification. Its appeal is largely intuitive, and little theoretical analysis is available. Making the method more effective is thus necessarily a matter of numerical experimentation.

The algorithm allows vectorization, and parallel computing could reduce computing time, especially in problems of a higher dimension. The method is also open to further improvement. The introduction of concepts such as long-term memory, diversification of search, and strategic oscillation could reduce the number of iterations without improvement. More sophisticated approaches to the neighborhood structure [different partitioning of the solution space or the use of techniques described (6)] may also be developed.

REFERENCES

1. L. C. W. Dixon and G. P. Szego, *Towards Global Optimization* (North-Holland, New York, 1975), vol. 1.
2. ———, *Towards Global Optimization* (North-Holland, New York, 1978), vol. 2.
3. S. Kirkpatrick, C. D. Gelatt Jr., M. P. Vecchi, *Science* **220**, 671 (1983).
4. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge Univ. Press, Cambridge, 1992).
5. P. M. Pardalos and J. B. Rosen, *Constrained Global Optimization: Algorithms and Applications* (Springer-Verlag, Berlin, 1987).
6. A. Törn and A. Žilniskas, *Global Optimization* (Springer-Verlag, Berlin, 1989).
7. C. Floudas and P. M. Pardalos, *Recent Advances in Global Optimization* (Princeton Univ. Press, Princeton, NJ, 1991).
8. K. D. Gibson and H. A. Scherega, in *Structure and Expression: From Proteins to Ribosomes*, M. H. Sarma and R. H. Sarma, Eds. (Adenine Press, Guilderland, NY, 1988), vol. 1, pp. 69–81.
9. R. Fletcher, *Practical Methods of Optimization* (Wiley, New York, 1980), vol. 1.
10. ———, *ibid.*, vol. 2.
11. P. J. M. van Laarhoven and E. H. L. Aarts, *Simulated Annealing: Theory and Applications* (Kluwer, Dordrecht, 1987).
12. D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning* (Addison Wesley, Reading, MA, 1989).
13. F. Glover, *ORSA J. Comput.* **1**, 190 (1989).
14. ———, *ibid.* **2**, 4 (1990).
15. A. Hertz and D. de Werra, *Computing* **39**, 345 (1987).
16. C. Frieden, A. Hertz, D. de Werra, *ibid.* **42**, 35 (1989).
17. A. Törn and A. Žilniskas, *Global Optimization* (Springer-Verlag, Berlin, 1989), p. 185.
18. R. S. Anderssen, in *Optimisation*, R. S. Anderssen, L. S. Jennings, D. M. Ryan, Eds. (Univ. of Queensland Press, St. Lucia, Australia, 1972), pp. 27–34.
19. Committee on the Next Decade of Operations Research (Condor), *Oper. Res.* **36** (1988).
20. A. Dekkers and E. Aarts, *Math. Program.* **50**, 367 (1991).

13 July 1994; accepted 23 November 1994

Simulation of Recent Global Temperature Trends

Nicholas E. Graham

Observations show that global average tropospheric temperatures have been rising during the past century, with the most recent portion of record showing a sharp rise since the mid-1970s. This study shows that the most recent portion of the global temperature record (1970 to 1992) can be closely reproduced by atmospheric models forced only with observed ocean surface temperatures. In agreement with a diverse suite of controversial observational evidence from the past 40 years, the upward trend in simulated tropospheric temperatures is caused by an enhancement of the tropical hydrologic cycle driven by increasing tropical ocean temperatures. Although it is possible that the observed behavior is due to natural climate variability, there is disquieting similarity between these model results, observed climate trends in recent decades, and the early expressions of the climatic response to increased atmospheric carbon dioxide in numerical simulations.

Observations indicate that global average surface air temperatures have been rising during the past century (*1*, *2*). Whether this trend is real and if so what processes are responsible are questions that have been the subject of much discussion, particularly because increased tropospheric temperature is one of the most consistent results from simulations of the effects of increasing concentrations of CO_2 and other greenhouse gases (*3*). The recent portion of the global tropospheric temperature record is characterized by a sharp rise beginning during the mid-1970s, with the signal of the El Niño–Southern Oscillation (ENSO) superimposed on the lower frequency changes (*4*). The results presented here show that this most recent portion of global air temperature record can be simulated closely by atmospheric general circulation models (GCMs) forced only with observed global ocean temperatures. Analyses show that both the ENSO time scale changes and the

Climate Research Division, Scripps Institution of Oceanography, La Jolla, CA 92093, USA.