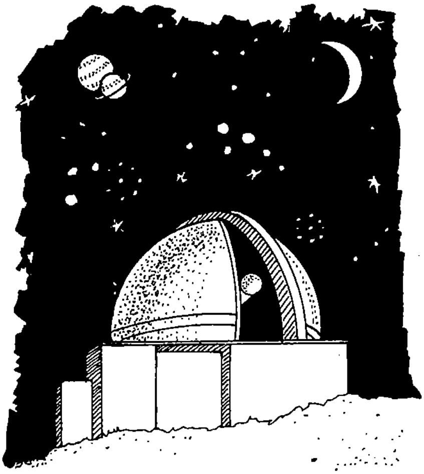

# Optimization and the Traveling Salesman Problem

**Charles A. Whitney** *Harvard University*

## INTRODUCTION

An astronomer at the controls of a telescope often wishes to make relatively short measurements of a long list of stars scattered across the sky. Moving the telescope from one star to another is a time-consuming process, so the astronomer tries to select the most efficient sequence of "visitations," avoiding unnecessary motion. This problem is particularly acute with a space telescope orbiting the Earth high above the atmosphere. The pointing of such a telescope is often controlled from the ground by small gas jets and the available fuel is severely limited. If the stars are haphazardly distributed around the sky, the sequence must be carefully chosen to conserve time and fuel.

This is a form of the "traveling salesman problem," and it is usually expressed as the search for the shortest closed route among a set of cities such that each city is visited just once. It is typical of a class of problems that can be stated briefly but that defy direct solution and can only be solved by a series of educated guesses. These guesses are often based on probabilistic calculations. This property distinguishes such problems from most of the problems described in this book, and it describes many practical problems in fields as different as geological exploration and electronics.

A direct approach to the solution would be to locate the cities on a map, measure all the possible routes, and select the shortest. But that is easier said than done, as we can see by counting the routes. If there are $N$ cities and the starting point is prescribed, there will be $N - 1$ choices for the first stop, $N - 2$ for the second, and so forth, making a total of $(N - 1) \times (N - 2) \times (N - 3) \ldots 2 \times 1$ possible routes back to the starting point. Even with as few as 8 cities, this leads to $7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 5{,}040$ possible routes to be examined. And the number increases rapidly with the number of cities. If we have two groups of 8 cities, making a total of 16, there would be more than a trillion routes! And 16 is not a particularly large number, so we evidently must give up the idea of a direct measurement of all possibilities.

Another example is the design of electronic circuits, where dozens of components are to be connected by wires. The designer needs a method of finding an efficient layout that will minimize the amount of wire required. In modern circuits the number of possible arrangements often exceeds a trillion, and the designer would not have time in the life of the universe to try all possible solutions while looking for the best.

## APPROACHING A SOLUTION OF THE TRAVELING SALESMAN PROBLEM

There is an aspect of this type of problem that gives us some hope of coming to a practical solution. We know intuitively that, in addition to the best solution, there will be many solutions that are nearly as good. Thus many of the good solutions, including the best, will be of nearly the same length. This means that we can be satisfied with a technique that brings us close to the optimum without worrying too much if we do not find the single best route. For problems of this type we can often assume that second best—or thousandth best—is good enough.

What more can we say about the possible solutions to the traveling salesman problem? Our intuition suggests—and experience verifies—that the traveler ought to complete the tour of each section of the map before moving to another section. As an extreme illustration, suppose the cities are arranged in two groups ($N/2$ in each group) separated by a distance of $G$ miles, as illustrated in Figure 1. A route that jumps back and forth between the two groups would clearly not be an efficient choice (see Figure 2). Its length would be roughly equal to the product, $N \times G$, and the salesman would do better to finish one group of $N/2$ cities before moving to the other, as in Figure 3. For example, suppose $D$ is the average distance between cities in each group, and the salesman starts from a selected city in the first group. A visit to all the remaining cities in the first group would involve a distance $D \times (N/2 - 1)$. Then he would move $G$ miles to a city in the other group, followed by $D \times (N/2 - 1)$ miles in visiting the remainder of the group. Finally, he would move $G$ miles to get back to the first group, for a total of $2G + D \times (N - 2)$ miles. Figures 2 and 3 compare the lengths of two such solutions, and they show that a little care can lead to a substantial shortening of the route.



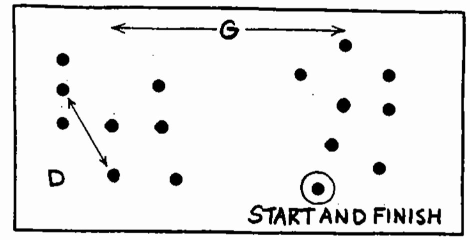


**Figure 1** *Schematic map of 16 cities. The traveling salesman problem is to find the shortest route from a specified city (circled) that passes just once through all the other cities. In this example the cities are divided into two groups separated by a distance of G miles. Within each group the mean separation is D miles.*

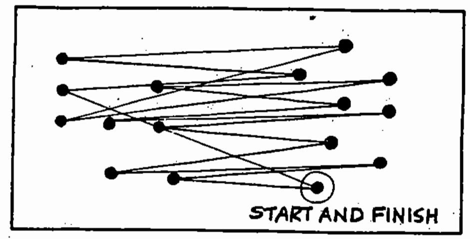


**Figure 2** *Example of a poor solution, in which the salesman moves back and forth between the two groups of cities. His route length is approximately $N \times G$ miles, where $N$ is the total number of cities. Figure 3 shows a better solution.*

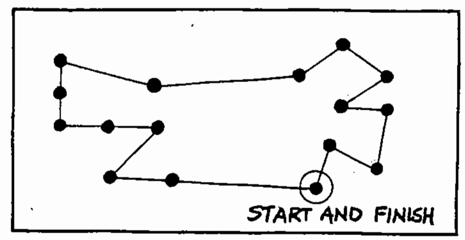


**Figure 3** *Example of a good solution, in which the salesman visits all the cities in one group before moving to the other. As shown in the text, his route length in this case is approximately $L' = 2 \times G + D \times (N - 2)$, which is less than the distance $L = N \times G$, for the route in Figure 2. As an example, suppose $N = 16$, $G = 8$ miles, and $D = 2$ miles. Then $L = 128$ miles for the route in Figure 2, and $L' = 2 \times 8 + 2 \times 14 = 44$ miles. Thus the length of the better route is about ⅓ that of the poorer route.*

So the salesman ought to cover each region thoroughly before moving to the next, but how does he search for a good route within each region?

Before tackling this problem, let's consider a more general type of problem that has many industrial and scientific applications. It may appear to be a new type, but it is simply a variation on the traveling salesman problem.

Consider the cost of fuel for the operation of a simplified manufacturing plant shown schematically in Figure 4. The plant buys electricity from a public utility to operate electric equipment such as power saws and an air conditioner. The plant also buys fuel oil to operate a generator that produces electricity and a certain amount of heat that can also be used for controlling the air temperature. Given the costs of electricity and oil, as well as the efficiency of the machines—which depends on the amount of effort put out by each machine—how should the plant manager allocate the money available for fuel, and how should the various types of fuel best be used within the plant? And if the plant needs more air conditioning capacity, what type of new machine ought to be brought in? In a realistically complex problem, the solution is far from obvious.

## GENERAL DESCRIPTION OF THE PROBLEM OF OPTIMIZATION

All of these problems have several features in common. First, for each, we can specify a *cost function* whose value we seek to reduce to a minimum. It may be the length of a route, or the cost of fuel, or the total time required to point a space telescope at a set of selected stars. Second, each problem is specified by a number of fixed parameters that are outside our control, such as the positions of the stars in the sky, or the fuel requirements of the factory machines and the manufacturing quotas set by management. These are *constraints* on the problem, and they limit the possible solutions. Third, there are the adjustable numbers whose choice constitutes the solution of the problem. For example, we adjust the sequence of stars until the time is minimized, or we adjust the purchases of fuel and the allocation of machine output in the factory to minimize the overall cost of operation.

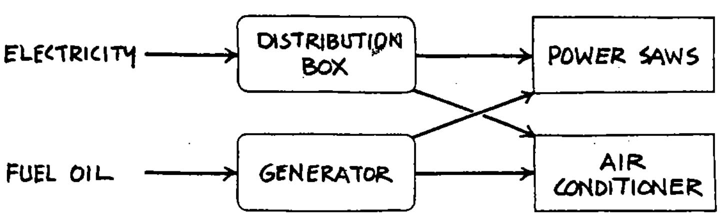


**Figure 4** *This diagram of a manufacturing plant that operates power saws and an air conditioner provides a simplified industrial example of the problems described in this essay. Electricity and gas have different costs and efficiencies, and the plant manager must decide on the most economical amounts of electricity and oil to buy. The answer will depend on deciding how much electricity to produce with the plant's own generator. In a real plant, the problem is often too complicated to be solved directly, and the manager must resort to random search techniques.*



With the advent of relatively inexpensive electronic computers, it has become feasible to carry out such searches by groping in the dark, so to speak. This process is very much like a randomized search for the deepest part of a lake.

To take a problem that arises in geological exploration, suppose we are prospecting for gas and are probing for the top of an underground gas dome, trapped above an oil field (Figure 5a). In order to get the largest quantity of gas, we wish to tap it at the highest part of the dome. Assume we have a device for measuring the height of the dome above the level surface of the oil at any point, and we look for the position that makes the height a maximum. For consistent terminology, we wish to express the problem in terms of minimizing a cost function, so let us define the depth of the dome below a convenient level surface as the cost function and seek to minimize it.

If we knew very little about the dome we would start at an arbitrary place. We would read the depth at the starting point, then move a short distance and again measure the depth. If it decreased, we would continue moving in that direction. If the depth increased, we would move in another direction. The selection of a new direction could be based on the local stratification or it might be a blind guess. In either case, we would continue this process until we came to a place where the depth increased in all directions. We would then know we were near the peak of the gas dome. All points within a small distance of the peak—say, a few meters—would have nearly the same depth because we can consider the top to be horizontal in the region of the peak. So we could adopt any of the positions in this region as the solution, as it would make no practical sense to insist on locating the peak to the nearest centimeter.

If the dome is actually a smooth spherical or cylindrical shape, we can be fairly sure we have come to the neighborhood of the true peak. But suppose the shape is more complex, as in Figure 5*b*. In that case, we may merely have found a localized peak. How can we avoid mistaking a local peak for the true top? The key is in a *probabilistic* approach that keeps us from getting caught at a false peak.

## DESCRIPTION OF A PROBABILISTIC SEARCH FOR THE OPTIMUM

In a probabilistic search we proceed by successive corrections, starting from a guessed solution and trying new, randomly generated solutions until we are satisfied. To see how this might work, let $C$ be the cost function to be minimized, say, the length of the route or the local depth of the gas dome, for a particular solution. We evaluate $C$ for the first guess and then construct a new solution—randomly or using whatever information we have. This is equivalent to moving to another position over the gas dome. We then reevaluate the cost function (depth), obtaining $C'$. If the new value is less than or equal to the previous, $C' \leq C$, we accept the new position as defining a next starting point. On the other hand, if the new depth is greater, $C' > C$, we decide whether to reject or accept the new position on the basis of a probabilistic calculation, as follows. We compute a random number (imitating the outcome of tossing a pair of dice, for example), and we restart from the old position if this number

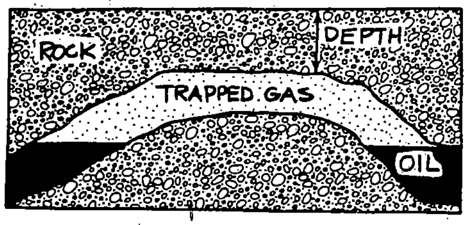


**Figure 5a** *Diagram of a hypothetical oil deposit with an overlying region of natural gas. The depth of the gas dome below the surface of the ground is indicated. When the gas is tapped off, the oil will rise in the dome, so the tap is to be placed at the highest peak of the gas dome to get all of the gas. The text describes a random search for the peak, where the depth is a minimum.*

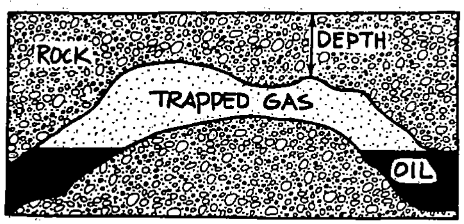


**Figure 5b** *Similar to Figure 5a, except the dome has a secondary peak that might be mistaken for the true peak. The search technique must be designed to avoid stopping at the secondary peak, or the gas in the higher peak will be lost.*

is less than some other number we have chosen previously. We repeat this process many times, moving from one point to the next.

Let us not worry, for the moment, about the details of the calculation that leads us to accept or reject a new solution because the details are not crucial to an understanding of the method. The essential point is that the probabilistic decision permits us to occasionally move to a new point even if it entails a small *increase* of $C$. This prevents our getting stuck in local minima before we reach the neighborhood of the absolute minimum. This is the way we avoid stopping at a secondary peak of the gas dome. We always permit steps that would momentarily take us to a deeper point of the gas dome because sometimes that is the only path to the true peak.

Despite the fact that the process is rather like groping in the dark, it has great power for two reasons. First, it searches among a small sample of possible solutions and picks out solutions that are approximately the best. This greatly reduces the search time. Second, it can be used to find a practical solution to any problem that can be expressed in terms of a cost function, $C$, and a set of adjustable parameters. As long as a computer is available, we need not care whether the numerical evaluation of the cost function is simple or complex. The sky is the limit, and this explains the usefulness of the random search for problems as diverse as manufacturing, geological exploration, and astronomy.

For those of you who are interested, I'll describe some of the technical details of the probabilistic decision to accept or reject each new solution. I will phrase it in terms of the randomized search for the peak of a gas dome.



After each new position has been selected and its depth, $C'$, has been evaluated, we divide the change of cost, $C' - C$ by a number, $T$, which has the same units as $C$. (The selection of the value of $T$ is a matter of experience. It must be suited to the problem.) We next compute a random number, $r$, in the range between zero and unity, $0 \leq r < 1$. We take its natural logarithm (which will always be a negative number) and compare the result with $-(C' - C)/T$. If

$$\text{logarithm } r < -(C' - C)/T,$$

we accept the new route; otherwise we reject it.

This particular formula is chosen because it is very permissive about small increases of $C$ while permitting only a very small number of large increases of $C$. The probability of accepting an increase of $C$ depends on the value of $T$. If $T$ is large, this formula permits large increases. This means that almost every new trial we make will be accepted as a new starting point. This is a good way to start a search because we can wander freely about the dome. As the computation progresses, it will sample the dome here and there and move among the various peaks. The rejection criterion is gradually adjusted by reducing the value of $T$, making the acceptance of an increase of $C$ less likely. After all, a geologist who has already sampled much of the gas dome will not want to move back to where the dome is lower. Establishing the schedule for reducing $T$ is a matter of experience; if $T$ is reduced too quickly, the process may get stuck at a false peak of the gas dome; if it is reduced too slowly, the geologist will wander indefinitely. Ideally, the reduction of $T$ will gradually restrict the search to the neighborhood of the true peak. The process is stopped when $T$ has become small and the search appears confined to a small region of the dome. The geologist can confirm that the true peak has been found by repeating the entire search and seeing whether it returns to the same region.

The effect of changing $T$ is illustrated in Figure 6*a* and *b*, where two hypothetical searches are traced. In the first case, which would be appropriate to the early stages of a search, the large value of $T$ permits probing much of the dome. When the search wanders back into the vicinity of the true peak, the value of $T$ is decreased, thus trapping the search in the best region.

## CONCLUSION

The random search method, despite its apparent blindness, has proven very powerful. The key to its success is twofold. It samples a small number of the total number of possible solutions, and it avoids being trapped by false solutions. A good solution for virtually any problem that can be expressed in terms of a cost function and a set of adjustable numbers can be found by this method. How close the derived solution comes to the best possible solution depends on the persistence of the solver, but in most practical problems, a good solution is entirely adequate.

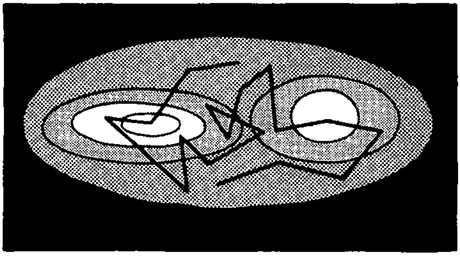

**Figure 6a** *Contour map of the roof of a gas dome with two peaks. The black region indicates the oil level; lighter shading indicates the higher roof, and the true peak is on the left. A random search with a large value of the parameter, T, is indicated by the segmented line. It covers a large portion of the dome, and it will not settle down in either peak until T is decreased.*

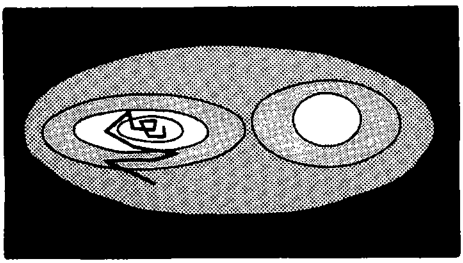

**Figure 6b** *Similar to Figure 6a. In this case, a small value of T was used and this confines the search to steps that move toward the nearest peak, so it can stop at the wrong peak, although in this case it did find the higher peak.*

## PROBLEMS

1. List several types of problems that might be amenable to the random search method. What are the characteristics of a problem that would require this type of approach?

2. Imagine you have a newspaper route and wish to find the most efficient delivery sequence. What are some of the ways you might construct trial routes? Suppose you live in a suburban area where the houses are strung along several roads with few cross streets. How would this affect the process by which you would construct trial routes?

## REFERENCE

S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi. 1983. "Optimization by Simulated Annealing." *Science* 220: 671–680.