the theory of ordered groups, and contains enough material for a one semester course or seminar.

PAUL CONRAD

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 84, Number 4, July 1978

Theory of optimal search, by Lawrence D. Stone, Mathematics in Science and Engineering, vol. 118, Academic Press, New York, 1975, xiv + 260 pp., $29.50.

That resources are limited and must be carefully allocated among competing ends, each in itself desirable, is a central fact of the world we live in. The analysis of the resource allocation problem for society as a whole has been a central concern of economic theory, while its study from more limited and more detailed perspectives has become perhaps the major focus of the field termed operations research or management science. Biological study, particularly the field of ecology and some aspects of evolutionary theory, has also put some emphasis on the resource allocation problems of living creatures; after all, Charles Darwin ascribed his notion of natural selection to the influence of the economist, Thomas R. Malthus, whose emphasis on the implications of resource limitations earned for economics the name of "the dismal science."

The book under review is a study of optimal resource allocation in a particular field, search for an object when the search process uses up scarce resources. This particular theory arose during World War II as the problem of locating enemy submarines and was studied by a group headed by the probability theorist, B. O. Koopman. Most of the subsequent interest has also been motivated by seeking submarines, including lost friendly ones. There is a considerable literature, to which the author has been a major contributor, and now we have a survey which is indeed admirable in scope and exposition.

Although the author concentrates on his particular area of resource allocation, more general problems are implied, and some of the theorems are widely applicable. Nevertheless, the search problem in its elementary forms has special features which enable stronger results to be obtained than are available generally.

The resource allocation problem with a single scarce resource can be stated as follows: Let there be a finite or denumerable set, J, of possible activities. Each can be operated at alternative levels indexed by a real number (possibly restricted to the integers, if the activity can be carried on only in discrete steps, or to some other subset of the reals). Let f be a mapping from J to the range of activity levels. If z is the activity level for the jth activity, let $c(j, z)$ be the amount of the scarce resource used in the jth activity. Hence, if f is the specification of activity levels, the total amount of the resource used is,

$$\sum_{j \in J} c(j, f(j)).$$

If the total amount of the resource is considered to be limited then we are restricting our choice of f to those for which the above expression equals some given constant, K.

To choose among these feasible allocations, the problem at hand suggests some criterion, to be maximized. Most generally, the criterion is a functional, $E[f]$. The elaboration of this theory, which is of course a problem in constrained optimization, to be handled by more or less sophisticated versions of the Lagrange multiplier method, occupies a vast literature in such fields as mathematical programming and control theory. In the case of search theory, however, the criterion has a special form, being additive. That is, we can write, $E[f] = \sum_{j \in J} e[j, f(j)]$. In other words, if the jth activity is operated at a level of z, then an input of amount $c(j, z)$ of the resource is required and a benefit of $e(j, z)$ is obtained.

Without any real loss of generality, it can be assumed that $c(j, z) \equiv z$. Then, if the functions $e(j, \cdot)$ are concave and increasing, usual Lagrange techniques show that optimal solutions are characterized by a parameter, λ, such that, for each j, $z = f(j)$ is chosen to maximize $e(j, z) - \lambda z$, with λ chosen so that the constraint on cost is satisfied.

In the above, J may be finite or denumerable. If instead we consider a continuous set of activities, say X, the maximand becomes $\int_X e[x, f(x)]\, dx$, and the cost constraint, $\int_X f(x)\, dx = K$. In this case, a somewhat stronger Lagrangian result is valid; without any concavity conditions, an optimal policy is characterized by a parameter, λ, such that, for almost all x, $z = f(x)$ is chosen to maximize $e(x, z) - \lambda z$. (That concavity conditions are unnecessary follows from Lyapunov's theorem that the range of a vector nonatomic additive set-function is convex.)

One more remark. It is easy to see that $f(j)$ [or $f(x)$] decreases as λ increases and therefore increases with K for each j or x. Hence, the optimal policies have what the author calls an *incremental* property: if an optimal policy has been found for K_1, and $K_1 < K_2$, then the optimal policy for K_2 under the constraint that effort for each j or x be at least equal to that for K_1 is in fact the unconstrained optimal policy.

The most usual single resources that might be constrained are money or time (which are frequently equivalent). If the resource is time, then the last remark has an obvious interpretation: if the problem first proposed is to maximize outcome for a given time and if then the time is extended, the maximizer will have no regrets.

A mapping $\phi(j, K)$ [or $\phi(x, K)$ in the continuous case] into the reals is *uniformly optimal* if, for each K, $\phi(\cdot, K)$ is optimal when costs are restricted to K (my notation differs slightly from the author's). The uniformly optimal policy is obtained, as just seen, by sweeping out the values of the Lagrange parameter and then matching them with corresponding values of K. If K is interpreted as total time available, then the policy is indeed uniformly optimal, in the sense that the allocation at any given moment, given by $df(j)/dK$, is the same for all K, for times not exceeding K. Different values of K only imply different stopping times.

Another optimality concept is that of *local optimality*. Having already allocated a total effort, K, imagine it incremented slightly and optimally allocate the increment. If this is done for each K, the resulting policy is called locally optimal, and, under the assumptions made, is the same as the uniformly optimal policy.

To apply these concepts specifically to the problem of search, let $p(j)$ be the *a priori* probability that the object sought for is in the jth cell of a search area (in the discrete case where J is finite or denumerable). Let $b(j, z)$ be the probability of detecting the object if it were in the jth cell when effort z is expended there; the function, $b(j, \cdot)$ is increasing. The probability of detection with an allocation, $f(j)$ is,

$$\sum_{j \in J} p(j) b[j, f(j)],$$

so that the general theory requires only the identification, $e(j, z) = p(j)b(j, z)$.

If the search area is considered continuous, then a similar identification can be made, with $p(x)$ the probability density over X, $b(x, z)$ the probability of identifying an object at x with effort z, and $e(x, z) = p(x)b(x, z)$.

It is natural in a search context to identify effort with time. In that case, the uniformly optimal policy has another interesting property: it minimizes mean time among all policies which continue to discovery. In fact, a more general property holds. Suppose we assign a finite value to the object. Then we might not wish to continue the search until discovery. We might then want to search for a while but then stop. It follows almost immediately from uniform optimality that for any such *search and stop* policy, the uniformly optimal policy should be followed so long as there is any search. However, the determination of the stopping time cannot be given a simple characterization in general.

Search and stop policies exemplify a broad class of important optimization problems, which may be called stochastic control or dynamic programming under uncertainty. The essential feature is that information is acquired during the process and used as the basis of further decision-making. This implies that earlier decisions must take into account the fact that the later decisions will be based on knowledge not now available. To take a simple example from economics: Let the decision today be the choice of a machine and that tomorrow the rate of production, determined by the demand which will be known tomorrow but not today. Some machines are highly specialized and efficient for one level of output but very costly to operate at other levels; other machines are less efficient at any given level of output than a specialized machine but have less variability over output levels. Clearly, if there is enough uncertainty at the initial date about the future demand, the flexible machine will be preferred.

In the search problem as stated, the informational feedback is of the simplest possible form; at any time t, the object has either been found or has not been found; the search terminates when the object is found. Because of this simple structure, the possibility of informational feedback does not alter the structure of the solution. But suppose we admit the possibility of false detection, which has so far been excluded; that is, the search process may affirm the finding of the object, and refutation requires a more detailed form of search. Then, at any moment the information includes the false targets found and the objects detected but not yet verified. Optimal solutions which use this information cannot be determined with the methods used (see Chapter VI); policies which are "optimal" neglecting the feedback can be found, but simple adaptive policies are shown to improve on them.

(If there is a significant criticism of the work under review, it is that insufficient attention is paid to algorithms which achieve or approximate optima in relatively complex situations, as opposed to problems which admit of elegant solutions.)

I have surveyed essentially the first six chapters of the book but have not done justice to the thoroughness and clarity of the exposition nor to the numerous and helpful examples. The remaining chapters deal with approximations and with moving targets, for which some interesting results are found, though not of the same generality as the earlier work.

KENNETH J. ARROW

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 84, Number 4, July 1978

Finite free resolutions, By D. G. Northcott, Cambridge Univ. Press, New York, xii + 271 pp., $29.50.

This book gives a beautifully self-contained treatment of the recent Buchsbaum-Eisenbud theory [4], [5] of finite free resolutions over a commutative ring with identity, as well as of a number of related topics (e.g. MacCrae's invariant [11]). There are two features in which the author's treatment differs from existing accounts of the subject: first, he confines himself almost entirely to elementary methods, avoiding Ext, Tor, and even exterior powers (we shall do likewise), and, second, he exploits a new notion of grade (or depth) in the non-Noetherian case which permits him to dispense entirely with the Noetherian restrictions on the ring. The very elementary form of the treatment enables the author to make accessible some fancy results from the homological theory of rings to readers with virtually no background in algebra.

Hilbert [7] gave the theory of finite free resolutions its initial impetus. Suppose that one is trying to understand a finitely generated module M over a Noetherian ring R (Noetherian means that every ideal is finitely generated, and implies that every submodule of a finitely generated module is finitely generated). To give generators $u_1, \ldots, u_{n_0}$ for M is essentially the same as to map a free module $F_0 = R^{n_0}$ onto M (the map then takes $(r_1, \ldots, r_{n_0})$ to $\Sigma_i r_i u_i$). To understand M, one simply needs to understand the kernel $\{(r_1, \ldots, r_{n_0}) \in R^{n_0}: \Sigma_i r_i u_i = 0\}$, call it $\operatorname{syz}^1 M$, of this map (of course it is not unique: it depends on the choice of generators). This kernel is called a *relation module* or *module of syzygies* for M. Note that $M \cong F_0/\operatorname{syz}^1 M$. But then, to understand $\operatorname{syz}^1 M$, it is entirely natural to choose, say, n_1 generators for $\operatorname{syz}^1 M$ (equivalently, to map $F_1 = R^{n_1}$ onto $\operatorname{syz}^1 M$) and so obtain a module of syzygies of the module of syzygies, denoted $\operatorname{syz}^2 M$. Of course, there is no reason to stop at this point, and so one can obtain a (usually infinite) sequence of modules of syzygies $\operatorname{syz}^i M$ each contained in a free module $F_{i-1} = R^{n_{i-1}}$. For each i we have a composite map $(F_i \twoheadrightarrow \operatorname{syz}^i M \hookrightarrow F_{i-1})$, call