ARAB REPUBLIC OF EGYPT

التخطيط القومى

THE INSTITUTE OF NATIONAL PLANNING

Memo. No. 1213

BRANCH-AND-BOUND REVISITED: A SURVEY OF BASIC CONCEPTS AND THEIR APPLICATIONS IN SCHEDULING

BY

SALAH E ELMAHGRABY
ALWALID N. ELSHAFEI

DEC. 1977

CAIRO
SALAH SALEM St.-NASR CITY

MODERN TRENDS IN LOGISTICS RESEARCH

Proceedings of a Conference Held at
The George Washington University

Edward S. Bres III
A. Charnes
W. W. Cooper
S. E. Elmaghraby
A. N. Elshafei
Richard L. Francis
Walter D. Gaddis
Donald P. Gaver
Murray A. Geisler
Donald Gross
Arnoldo C. Hax
Fred Kornet, Jr.
Arthur I. Mendolia
Frank Proschan
David A. Schrady
Jeremy F. Shapiro
Ronald W. Shephard
William W. Snavely
Gerald L. Thompson
S. Zacks

Edited by

W. H. Marlow

The MIT Press
Cambridge, Massachusetts, and London, England

CONTRIBUTORS

Edward S. Bres III is a Research Associate, Center for Cybernetic Studies, University of Texas at Austin.

A. Charnes is the University of Texas System Professor, Jesse H. Jones Professor of Biomathematics and Management Science, and Director, Center for Cybernetic Studies, University of Texas at Austin.

W. W. Cooper is University Professor of Public Policy and Management Science, School of Urban and Public Affairs, Carnegie-Mellon University.

S. E. Elmaghraby is University Professor of Operations Research and Industrial Engineering, and Director, Operations Research Program, North Carolina State University.

A. N. Elshafei is Expert, The Institute of National Planning, Nasr City, Cairo, Egypt.

Richard L. Francis is Professor, Department of Industrial and Systems Engineering, University of Florida.

Walter D. Gaddis served as Deputy Chief of Naval Operations (Logistics) from April 11, 1973 to August 1, 1975.

Donald P. Gaver is Professor and Associate Chairman for Research, Department of Operations Research and Administrative Sciences, United States Naval Postgraduate School.

Murray A. Geisler is former Director of Logistics Studies, Rand Corporation, Santa Monica, California.

Donald Gross is Professor, Department of Operations Research, George Washington University.

Arnoldo C. Hax is Associate Professor of Management Science, Alfred P. Sloan School of Management, Massachusetts Institute of Technology.

Fred Kornet, Jr. served as Deputy Chief of Staff for Logistics, Department of the Army, from January 2, 1973 to August 31, 1975.

W. H. Marlow is Professor and Chairman, Department of Operations Research, Principal Investigator of Program in Logistics, and Director, Institute for Management Science and Engineering, George Washington University.

Arthur I. Mendolia served as Assistant Secretary of Defense (Installations and Logistics) from June 21, 1973 to March 31, 1975.

Frank Proschan is Professor, Department of Statistics, Florida State University.

David A. Schrady is Chairman, Department of Operations Research and Administrative Sciences, United States Naval Postgraduate School.

Jeremy F. Shapiro is Professor of Operations Research and Management, Alfred P. Sloan School of Management, Massachusetts Institute of Technology.

Ronald W. Shephard is Professor of Engineering Science and Chairman, Department of Industrial Engineering and Operations Research, University of California, Berkeley.

William W. Snavely served as Deputy Chief of Staff for Systems and Logistics, Department of the Air Force, from February 1, 1973 through August 31, 1975.

Gerald L. Thompson is Professor of Applied Mathematics and Industrial Administration, Carnegie-Mellon University.

S. Zacks is Professor and Chairman, Department of Mathematics and Statistics, Case Western Reserve University, and Consultant, Program in Logistics, George Washington University.

Chapter 8

BRANCH-AND-BOUND REVISITED: A SURVEY OF BASIC CONCEPTS AND THEIR APPLICATIONS IN SCHEDULING*

S. E. Elmaghraby and A. N. Elshafei**
North Carolina State University

8.1 Preliminaries

The term "branch-and-bound" (B&B) has increasingly become a household term among students and researchers in the field of scheduling and sequencing. In this chapter we shall take a fresh look at this approach and assess its content, utility, and potential. In delineating the subject matter of our discussion, perhaps it is equally valid to emphasize that which is not among our aims. This chapter is not a comprehensive survey of B&B concepts and applications. Several survey articles that have appeared in recent years serve that function adequately, if not superbly; see, for example, References [8.1, .4, .21, .36, .40]. Nor does this paper aspire to be a comparative evaluation of the very many B&B approaches that have been proposed in the open literature to solve one scheduling problem or another. For examples of such studies, the reader is referred to the papers of Ashour [8.2], Ashour and Quraishi [8.3], Davis [8.8], and Kan [8.34], among others.

What we do wish to present is an inventory of the basic concepts underlying the "theory" of B&B; we wish in fact to establish that such theory exists and to illustrate these basic concepts by examples from the field of scheduling and sequencing. In this we are

---

*The preparation of this chapter was partially supported by the Office of Naval Research under Contract N00014-70-A-0120-0002, by the National Science Foundation under Grant P1K1470-000, and by the Army Research Office-Durham under Contract DA-ARO-D-31-124-72-G106, with North Carolina State University.

**Now at the Institute of National Planning, Cairo, Egypt

motivated by two objectives. The first is to summarize, in what we hope is a convenient place, the multitude of concepts that have emerged over the past few years. We hope that such a summary will provide a handy reference and basic understanding to student and researcher alike. The second is to help the profession assess the current and future potential of this approach. In this respect, one may compare B&B as a problem solving approach, to simulation which is another, and by now a very popular problem-solving approach. One may then ask fundamental questions similar, but not necessarily identical, to those asked in the study of simulation. For instance, in Monte Carlo simulation one often raises the question of variance-minimizing techniques. In B&B one may ask questions relative to the rate of convergence to the optimum.

As much as possible we shall draw our examples from the field of scheduling and sequencing. However, since problems of scheduling (and sequencing) are almost universally modeled as integer or mixed programming problems (linear or nonlinear), we shall feel free to illustrate some concepts with reference only to the integer (or mixed) program, without the need to motivate the model by the scenario of the scheduling problem. Ordinarily, we shall be dealing with integer linear problems (ILP) and, in particular, with 0,1 ILPs. As is well known, an ILP can be translated into a 0,1 ILP by the simple binary expansion of the variables. In a couple of instances, we could not find examples from scheduling and, to the best of our knowledge, none exist that use a particular concept. Then we took the liberty to illustrate by examples from other fields of application, such as location-allocation. We do not feel particularly apologetic about taking such liberty since these problems are themselves modeled as integer (linear or nonlinear) programs. Such models provide the link to problems in scheduling.

In the sequel we shall be talking about "partial solutions" and "completions." The term "partial solution" is actually a misnomer, since it refers to something that provides no solution whatsoever to the original problem. For instance, a schedule of a subset of the jobs, or a series of cities visited by the salesman in the traveling salesman problem (TSP), are referred to as partial solutions, yet they provide no "solution" to the problems posed, which are: a complete schedule of all jobs in the first case, and a complete tour over all cities in the second case. The reader will hopefully bear with this misuse of language. By the completion of a partial solution we mean the specification of the values of the remaining variables so that their union with the partial solution yields a point in the original solution space. A partial solution is said to be fathomed if one of the two following conditions is satisfied.

(1) We determine that its best feasible completion is better (yields a better objective value) than the best feasible solution known to date (assuming one is in hand).

(2) We determine that the partial solution has no feasible completion better than the incumbent (this includes infeasibility, which is translated into infinite penalty).

The concept of fathoming is illustrated in Example 8.3.

8.2 Fundamentals

The approach of B&B is basically a heuristic tree search in which the space of feasible solutions, which may contain a very large (or denumerable) number of points, is systematically searched for the optimum. According to Mitten and Warburton [8.42], "the search proceeds iteratively by alternately applying two operations: subset formation and subset elimination. In the former, new subsets of alternatives are formed, while in the latter some subsets of alternatives may be eliminated from further consideration. The procedure terminates when a collection recognized to contain only optimal solutions is reached." The search has two guiding principles: first, that every point in the space is enumerated either explicitly or implicitly, and, second, that the minimum number of points be explicitly enumerated. (We view B&B as an approach for implicit enumeration, though we concede that, mainly due to historical coincidences, the label "implicit enumeration" has been applied to approaches that need not employ the "bounding" feature of B&B.)

The implicit enumeration of feasible points is accomplished through dominance (which may or may not employ bounding) and feasibility considerations. Each of these concepts will be discussed in greater detail below, but first we give a laconic description of them to afford the uninitiated reader a general grasp of the subject. The basic idea in B&B is to divide the feasible space, denoted by S, into subsets $S_1, S_2, \ldots, S_k$ which may or may not be mutually disjoint. Assuming that the optimum falls in subset S_k, a bound on its value is determined: an upper bound (u.b.) in the case of maximization, and a lower bound (ℓ.b.) in the case of minimization or, better still, both an upper and a lower bound in either case. Based on such bounds two actions may take place: (i) a particular subspace S_k is selected for more intensive search by further partitioning into its subsets (this is the branching, or "formation" function); (ii) some feasible points (subspaces) are declared "noncandidates" for the optimum, and thus are eliminated from further considerations. This latter idea is one of "dominance" since it is based on the determination that any element of a particular subset S_i is better (or worse), in the sense of the criterion function, than any element in another subset S_j. Then indeed we may declare the points in S_j (or in S_i) as noncandidates for the optimum and eliminate them from further analysis.

While dominance may be established on the basis of the bounds evaluated on subsets S_k, it is also true that dominance can be established independent of any bounding considerations. In some circles (especially in the scheduling literature) these are referred to as "elimination" procedures. The final result is the same, namely, it establishes that certain subspaces cannot contain the optimum because they are dominated by other subsets. A similar idea lies behind the feasibility considerations. They arise because in the majority of cases one is forced to hypothesize a rather "rich" original space S. At some stage of analysis, if it can be established that certain subsets of S are in fact infeasible (in the sense of violating some constraint of the problem), then indeed such subsets can also be eliminated from further study.

Heuristics enter the tree search in all three basic phases of the approach: in the definition of the partitioning procedure, in the calculation of the bounds, and in the philosophy of searching the tree. But we wish to draw the reader's attention to the following important and rather crucial distinction: the formal structure of B&B admits the use of heuristics (as does the simplex algorithm of linear programming). However, these are "reliable heuristics" in the sense that if they run to completion, the optimum will be achieved. Furthermore, if the procedure is terminated before it has achieved the optimum, it yields a bound on the error committed. (This is in sharp contrast to "heuristic problem-solving procedures" which lay no claim to either optimality or to measuring the error committed at premature abortion.)

A more formal definition of the B&B procedure was advanced by Mitten [8.39] in 1970 which was expanded upon in later work in 1973 by Mitten [8.40], and Mitten and Warburton [8.42]. Mitten defines the operations of "branching," "bounding," and "branch-and-bounding" in terms of set functions. The necessary properties of each function were given in terms of operator and operands, which map all the known concepts of B&B into topological domains. He establishes the relations between the B&B recursive function and the set of optimal feasible solutions by postulating various analytic and topological conditions such as continuity, completeness, and compactness. In the case of finite solution space, the convergence of the B&B recursive function is easily seen. However, in denumerable or nondenumerable spaces, Mitten demonstrates that if the B&B recursive function is a contraction mapping in a complete metric space with appropriately defined elements, then fixed-point theorems could be invoked to establish convergence.

To gain more insight into Mitten's construction, we assume that it is desired to solve the problem: maximize $f(x)$ for $x \in X$. (For example, X may be the integer feasible points in an ILP.) Typically, B&B

proceeds by searching the space $T \supseteq X$ for the set $\sigma^* \triangleq \{x \in X: f(x) = f^*\}$ of optimal solutions, where $f^* = \sup_{x \in X} f(x)$ and where $\triangleq$ means "is defined to equal." The search proceeds by examining subsets $\sigma \in X$, and collections of such subsets. Let S denote the family of all possible collections of subsets that could be encountered by a given B&B procedure. As shorthand notation, let $\bigcup(s)$ denote a subset of X comprised of the elements in $\bigcup_{\sigma \in s}\{\sigma\}$; that is

$\bigcup(s) \triangleq \bigcup_{\sigma \in s}\{\sigma\}$ where $s \in S$.

As mentioned above, alternative possibilities in B&B are considered in sets rather than one at a time. Furthermore, B&B examines successively smaller and smaller subsets of X (the subset formation operation), always eliminating those subsets that can be shown not to contain an optimal solution (the elimination operation). It is assumed that once sets are "small enough" in some sense, then there is a procedure available for distinguishing the optimal solutions from the nonoptimal solutions, the so-called fathoming procedure. Therefore, let S^- denote the set of fathomable collections $\{s\}$; here s is a fathomable collection iff $\sigma \in s$ satisfies $\sigma \subset \sigma^*$ or $\sigma \cap \sigma^* = \phi$. We assume that a procedure is available for separating one from the other. That is, $s \in S^-$ iff the following hold.

(a) $s = s_1 \cup s_2$ with $\sigma \subseteq \sigma^*$ for every $\sigma \in s_1$ and $\sigma \cap \sigma^* = \phi$ for every $\sigma \in s_2$

(b) There is a means available for forming the collections s_1 and s_2 .

As a minimum requirement, we insist that any collection of singleton sets (sets containing one point of X each) is fathomable, since such sets cannot be subdivided. One may now state the objectives as: find a collection $s^* \in S$ such that $\bigcup(s^*) \subseteq \sigma^*$ and

$\bigcup(s^*) = \phi$ only if $\sigma^* = \phi$. Mitten defines the B&B procedure in terms of the set operations called branching, upper bounding, lower bounding, and, finally, branch-and-bounding. Let S_F (for formation) and S_E (for elimination) be two subfamilies of S. Branching may be defined as a function $F: S_F \to S_E$ such that for each $s \in S_F$ the following hold.

(a) $F(s) = \bigcup_{\sigma \epsilon s} \{d(\sigma)\}$, where $d(\sigma)$ is either σ or a collection of proper subsets of σ whose union is σ

(b) $F(s) = s$ iff $s \in S^-$

In words, this latter condition (a) states that each σ in s either remains unchanged under $F(s)$ or is broken up into a collection of proper subsets. This is illustrated in Figure 8.1, in which $s = \{\sigma_1, \sigma_2, \sigma_3\}$; $d(\sigma_2) = \sigma_2$; $d(\sigma_3) = \sigma_3$, but σ_1 has been "broken up" into four (disjoint) subsets. Clearly, $\bigcup_j \sigma_{1j} = \sigma_1$.
Upper bounding is a real-valued function $u: \bigcup(S_E) \to \underline{\underline{R}}$ with the following properties.

(a) $u(\sigma) \geq f(x)$ for all $x \in \sigma \in \bigcup(S_E)$

(b) $u(\sigma) \geq u(\sigma_0)$ if $\sigma_0 \subseteq \sigma$; σ_0 , $\sigma \in \bigcup(S_E)$

(c) $u(\{x\}) = f(x)$, $x \in \sigma$

These latter conditions (a) and (b) follow from the common concepts of upper bounds and set inclusion. Condition (c) ensures that the upper bound on singleton subsets is the "value" of that point under the mapping f . Lower bounding is a real-valued function $\ell: S_E \to \underline{\underline{R}}$ such that the following hold for any $s \in S_E$.

(a) $\ell(s) \leq f^*$

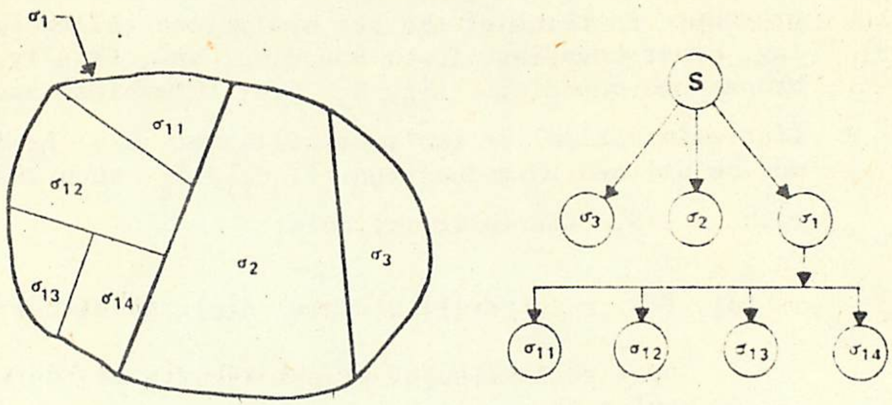

Figure 8.1 - Partitioning and branching.

(b) $\ell(s) \leq \ell(F(s))$

(c) $\ell(s) \geq f(x)$ for any $x \in s$

(d) If $s' \subset s$ is such that, for every $\sigma' \in s'$ either $u(\sigma') = -\infty$ or $u(\sigma') < \ell(s)$, then $\ell(s') = \ell(s)$.

These latter conditions (a) and (b) follow from the common concepts of lower bound and set division into subsets. Condition (c) ensures that the lower bound of a singleton subset is tight. Condition (d) guarantees that infeasible sets $(u(\sigma') = -\infty)$ or dominated sets $(u(\sigma') < \ell(s))$ cannot affect the value of the lower bound $\ell(s)$.

In Mitten's view, the bounding operation is an elimination operation through infeasibility and dominance. He defines it as a function $E: S_E \rightarrow S_F$ defined for $s \in S_E$ by

$E(s) = s - \{\sigma \varepsilon s: u(\sigma) = -\infty \text{ or } u(\sigma) < \ell(s)\}$

The strict inequality in the above statement implies strong bounding, since all optimal solutions in σ^* are retained. If an inequality is substituted, the resulting bounding operation is said to be weak since we then ensure that at least one element of σ^* will be retained. Finally, the B&B recursive operation is a function $G: S_F \to S_F$ defined by $G(s) = E(F(s))$. In other words, the successive formation and elimination of subsets is the heart of the procedure, hopefully leading to an optimum without the need to enumerate all singleton subsets.

If X contains finitely many points, it can be shown that $s^n = G(s^{n-1})$, for some finite $n > 1$, is an element of the fathomable set S^- , so that the procedure will terminate in a finite number of iterations. In the case X is not finite, Mitten shows that G will not "cycle" provided that each collection in S_F and S_E contains only finitely many sets. (Cycling means that there exists an m such that $G^m(s) = s$ and $s \varepsilon S_F - S^-$.) Note that even though a procedure may never cycle, it may not terminate in a finite number of steps. With this formal structure established, Mitten proceeds to illustrate his concepts by two examples: ILP and sequential unimodal search. This latter illustration is interesting since it claims to be the following.

(i) An example in which neither the procedure nor the sets involved are finite.

(ii) The only currently known application of B&B methods employing a branching rule that can be demonstrated to be optimal (the Fibonacci search).

This led Mitten to the following two conclusions. First, that the existence of an optimal branching rule for the sequential unimodal search suggests some interesting avenues of investigation in other areas of application. Second, that since attempts to extend the sequential search method to higher dimensions $\underline{\underline{R}}^n (n>1)$ have been notably unsuccessful, perhaps a fresh attack on the problem in $\underline{\underline{R}}^n$ via B&B would provide a new perspective. We wish only to remark that viewing sequential unimodal search as an application of B&B may raise some eyebrows, since none of the concepts usually associated with B&B are present in the standard search pattern, including the optimal pattern. The viewing is justified, however, if one sticks to the formal definition of B&B's search as composed of set formation and set elimination, both of which are indeed present in sequential unimodal search.

8.3 Branching

Branching proceeds by dividing the solution space into subspaces, which are themselves divided into subspaces, and so forth, until subsets containing exactly one point each are reached. The graphic representation is a tree, the search tree, whose numbering runs opposite to the set content. Thus, S_0 is the empty set ϕ which represents in fact the whole space S before any division has taken place. A terminal node of the tree, S_M, M large, contains a complete solution X which represents in fact a singleton set. Intermediate nodes of the search tree generally represent partial solutions generically represented by S_k. Hereafter we use the terms "branching" and "dividing the solution space" synonymously. The choice of the node from which to branch is basically a decision related to the philosophy of searching the tree, which is open to the use of heuristics.

Basically, there are two extreme philosophies with innumerable intermediate variations. On the one end of the spectrum there are heuristics (for example, branch from the node with the smallest lower bound) that favor the nodes higher up the tree. In this case, the construction of the search tree will proceed "horizontally"; this is the so-called jump-tracking (or "flooding")