SOME LINEAR PROGRAMMING ASPECTS OF COMBINATORICS

by

V. Chvátal

STAN-CS-75-505
SEPTEMBER 1975

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

Some Linear Programming Aspects of Combinatorics

by

V. Chvátal

Abstract

This is the text of a lecture given at the Conference on Algebraic Aspects of Combinatorics at the University of Toronto in January 1975. The lecture was expository, aimed at an audience with no previous knowledge of linear programming.

This research was supported in part by National Science Foundation grant DCR72-03752 A02 and by the Office of Naval Research contract NR 044-402.

Some Linear Programming Aspects of Combinatorics

V. Chvátal

Abstract

This is the text of a lecture given at the Conference on Algebraic Aspects of Combinatorics at the University of Toronto in January 1975. The lecture was expository, aimed at an audience with no previous knowledge of linear programming.

1. Introduction: Two Examples

In 1928, Sperner [33] answered the following question Let F be a family of distinct subsets of $\{1,2,\ldots,n\}$ such that

$$S,T \in F \Rightarrow S \not\subset T \; ; \tag{1.1}$$

how large can $|F|$ be? Sperner proved that

$$|F| \leq \binom{n}{\lfloor n/2 \rfloor} \; . \tag{1.2}$$

(To see that this is the best possible result, consider all the subsets S of $\{1,2,\ldots,n\}$ with $|S| = \lfloor n/2 \rfloor$ ●) In 1966, Lubell [28] gave a very elegant proof of this result; slightly recast, Lubell's argument goes as follows. Let us denote by A the family of all 2^n subsets of $\{1,2,\ldots,n\}$; let us call a family F feasible if it satisfies (1.1). With each family F, feasible or not, associate the vector $(x_S : S \in A)$ defined by

$$x_S = \begin{cases} 1 & \text{if } S \in F \; , \\ 0 & \text{if } S \notin F \; . \end{cases}$$

Thus obviously

$$|F| = \sum_{S \in A} x_S \; . \tag{1.3}$$

A family of sets $T_0, T_1, \ldots$ [illegible] with

$$\emptyset = T_0 \subset T_1 \subset \cdots \subset T_n = \{1,2,\ldots,n\}$$

will be called a chain. Clearly, there are n! distinct chains, each $S \in A$ is included in $|S|!(n - |S|)!$ of them. Furthermore, F is feasible if and only if

$$\sum_{S \in C} x_S \leq 1 \qquad \text{for every chain } C \text{ .} \tag{1.4}$$

The sum of all these n! inequalities (1.4) reads

$$\sum_{S \in A} |S|!(n - |S|)! x_S \leq n!$$

or, equivalently,

$$\sum_{S \in A} \frac{1}{\binom{n}{|S|}} x_S \leq 1 \text{ .}$$

Since every x_S is nonnegative and every $\binom{n}{|S|} \leq \binom{n}{\lfloor n/2 \rfloor}$, we have

$$\sum_{S \in A} \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} x_S \leq \sum_{S \in A} \frac{1}{\binom{n}{|S|}} x_S \text{ .}$$

Thus (1.4) implies

$$\sum x_S \leq \binom{n}{\lfloor n/2 \rfloor}$$

which, by virtue of (1.3), is the desired result.

Our second example goes back to the thirties when Erdös, Ko, and Rado [15] answered the following question. Let F be a family of k-element subsets of $\{1,2,\ldots,n\}$ such that

$$S,T \in F \Rightarrow S \cap T \neq \emptyset \text{ ;} \tag{1.5}$$

how large can $|F|$ be? Erdös, Ko and Rado proved that (in the nontrivial case $n \geq 2k$)

$$|F| \leq \binom{n-1}{k-1} \text{ .}$$

(To see that this is the best possible result, consider all the k-element subsets S of $\{1,2,\ldots,n\}$ with $1 \in S$.) In 1972, Katona [25] gave a very elegant proof of this result; slightly recast, Katona's argument goes as follows. Let us denote by A the family of all $\binom{n}{k}$ subsets of $\{1,2,\ldots,n)$ having k elements; let us call a family F feasible if it satisfies (1.5). For simplicity, let us assume (unlike Katona) that k divides n and let us set $m = n/k$. A family of pairwise disjoint sets $T_1, T_2, \ldots, T_m \in A$ will be called a partition. Clearly, there are exactly

$$\binom{n}{k}\binom{n-k}{k}\binom{n-2k}{k} \cdot \ldots \binom{k}{k} = \frac{n!}{(k!)^m}$$

ordered partitions; every $S \in A$ is included in

$$m \frac{(n-k)!}{(k!)^{m-1}}$$

of them. Furthermore, F is feasible if and only if

$$\sum_{S \in P} x_S \leq 1 \quad \text{for every partition } P .$$

The sum of all these inequalities reads

$$m \frac{(n-k)!}{(k!)^{m-1}} \sum_{S \in A} x_S \leq \frac{n!}{(k!)^m}$$

or, equivalently,

$$\sum_{S \in A} x_S \leq \frac{1}{m}\binom{n}{k} = \binom{n-1}{k-1}$$

which is the desired result.

In each of our two examples, the proof came out rather effortlessly. Was it just plain luck, one may wonder, or are we actually onto something? The answer to this ill-posed question is ambiguous. We were lucky indeed: proofs like that are not to be found for every combinatorial theorem. At the same time, however, we are onto something. We are onto the duality theorem of linear programming.

2. The Duality Theorem of Linear Programming

In each of the two introductory examples, we have argued that a certain set of linear inequalities (corresponding to the assumptions of our theorem) implies another linear inequality (corresponding to the desired conclusion). In general, we shall say that a set

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \qquad (1 \leq i \leq m) \tag{2.1}$$

of linear inequalities <u>implies</u> an inequality

$$\sum_{j=1}^{n} c_j x_j \leq d \tag{2.2}$$

if, and only if,

(i) there is at least one solution of (2.1) and

(ii) every solution of (2.1) satisfies (2.2).

When $y_1, y_2, \ldots, y_m$ are nonnegative reals, we call the inequality

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} y_i a_{ij} \right) x_j \leq \sum_{i=1}^{m} y_i b_i$$

a <u>linear combination</u> of (2.1) with <u>multipliers</u> $y_1, y_2, \ldots, y_m$.

Furthermore, an inequality

$$\sum_{j=1}^{n} a_j x_j \leq b$$

is called a <u>combination</u> of (2.1) if, for some b^* with $b^* \leq b$, the inequality $\Sigma a_j x_j \leq b^*$ is a linear combination of (2.1). Clearly, if (2.1) has at least one solution then it implies each of its combinations.

THE DUALITY THEOREM (first version). If (2.1) implies (2.2) then (2.2) is a combination of (2.X).

Customarily, the duality theorem is stated in a slightly different form. This form arises in the study of <u>linear programming problems</u> (or LP problems for short) such as

$$\left.\begin{array}{ll} \text{maximize } \sum_{j=1}^{n} c_j x_j \text{ subject to the constraints} & \\ \sum_{j=1}^{n} a_{ij} x_j \leq b_i & (1 \leq i \leq m) \\ x_j \geq 0 & (1 \leq j \leq n) \end{array}\right\} \tag{2.3}$$

With (2.3) one associates another LP problem, called the dual of (2.3):

$$\left.\begin{array}{lr} \text{minimize } \sum_{i=1}^{m} b_i y_i \text{ subject to the constraints} & \\ \sum_{i=1}^{m} a_{ij} y_i \geq c_j & (1 \leq j \leq n) \\ y_i \geq 0 & (1 \leq i \leq m) \quad . \end{array}\right\} \quad (2.4)$$

A solution to the constraints in (2.3), resp. (2.4), is called a feasible solution of (2.3), resp. (2.4). A feasible solution which maximizes $\sum c_j x_j$, resp. minimizes $\sum b_i y_i$, is called an optimal feasible solution of (2.3), resp. (2.4). Note that for every feasible solution $x_1^*, x_2^*, \ldots, x_n^*$ of (2.3) and for every feasible solution $y_1^*, y_2^*, \ldots, y_m^*$ of (2.4), we have

$$\sum_j c_j x_j^* \leq \sum_{i,j} a_{ij} x_j^* y_i^* \leq \sum_i b_i y_i^* \quad . \qquad (2.5)$$

THE DUALITY THEOREM (second version). If (2.3) has an optimal feasible solution $x_1^*, x_2^*, \ldots, x_n^*$ then (2.4) has an optimal feasible solution $y_1^*, y_2^*, \ldots, y_m^*$ and

$$\sum_{j=1}^{n} c_j x_j^* = \sum_{i=1}^{m} b_i y_i^* \quad .$$

It is easy to see that the first version implies the second. Indeed, let $x_1^*, x_2^*, \ldots, x_n^*$ be an optimal feasible solution of (2.3); set $d^* = \sum c_j x_j^*$. Then the inequalities

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \qquad (1 \leq i \leq m)$$

$$-x_j \leq 0 \qquad (1 \leq j \leq n)$$

imply the inequality

$$\sum_{j=1}^{n} c_j x_j \leq d^* .$$

By the first version of the duality theorem, there are nonnegative reals $y_1^*, y_2^*, \ldots, y_{m+n}^*$ such that

$$\sum_{i=1}^{m} y_i^* a_{ij} - y_{m+j}^* = c_j \qquad (1 \leq j \leq n)$$

and

$$\sum_{i=1}^{m} y_i^* b_i \leq d^* .$$

Thus $y_1^*, y_2^*, \ldots, y_m^*$ is a feasible solution of (2.4) with

$$\sum_{i=1}^{m} b_i y_i^* \leq \sum_{j=1}^{n} c_j x_j^* .$$

BY (2.5), the last inequality must hold with sign of equality and $y_1^*, y_2^*, \ldots, y_m^*$ is an optimal feasible solution of (2.4).

To deduce the first version from the second, let us assume that (2.1) implies (2.2) and let us consider the following LP problem:

$$\text{maximize} \quad \sum_{j=1}^{n} c_j (u_j - v_j)$$

$$\text{subject to} \quad \sum_{j=1}^{n} a_{ij} (u_j - v_j) \leq b_i \qquad (1 \leq i \leq m)$$

$$u_j \geq 0 , \quad v_j \geq 0 \qquad (1 \leq j \leq n) .$$

For every real x_j, we may write $x_j = u_j - v_j$ with $u_j, v_j > 0$. Therefore our problem has an optimum feasible solution u_j^*, v_j^* $(1 < j \leq n)$; in fact, $\sum c_j (u_j^* - v_j^*) \leq d$. By the second version of the duality theorem, there are nonnegative reals $y_1, y_2, \ldots, y_m$ with

$$\sum_{i=1}^{m} a_{ij} y_i \geq c_j \quad ,$$

$$\sum_{i=1}^{m} (-a_{ij}) y_i \geq -c_j \quad ,$$

$$\sum_{i=1}^{m} b_i y_i = \sum_{j=1}^{n} c_j (u_j^* - v_j^*) \quad .$$

Hence (2.2) is a combination of (2.1).

Finally, we shall restate the duality theorem in yet another form. The set (2.1) of linear inequalities is called <u>inconsistent</u> if there are nonnegative reals y_i $(1 \leq i \leq m)$ such that

$$\sum_{i=1}^{m} a_{ij} y_i = 0 \qquad (1 \leq j \leq n) \quad ,$$

$$\sum_{i=1}^{m} b_i y_i < 0 \quad .$$

Trivially, an inconsistent set (2.1) is unsolvable; again, the converse is given by the duality theorem.

THE DUALITY THEOREM (third version). The set (2.1) is unsolvable if and only if it is inconsistent.

This version follows easily from the first version. Indeed, assume that (2.1) is unsolvable and let k be the largest subscript such that the set

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \qquad (1 \leq i < k) \tag{2.6}$$

is solvable. The set of all the solutions of (2.6) is a closed convex, and possibly unbounded, subset of R^n ; the assignment

$$(x_1, x_2, \ldots, x_n) \mapsto \sum a_{kj} x_j$$

maps this set onto a closed interval I with

$$z \le b_k \Rightarrow z \notin I \quad .$$

Hence there is some d with $d > b_k$ such that (2.6) implies

$$\sum_{j=1}^{n} (-a_{kj}) x_j \le -d \quad .$$

By the first version of the duality theorem, there are nonnegative reals y_i $(1 \le i < k)$ such that

$$\sum_{i=1}^{k-1} a_{ij} y_i = -a_{kj} \qquad (1 \le j \le n) ,$$

$$\sum_{i=1}^{k-1} b_i y_i \le -d \quad .$$

Setting $y_k = 1$ (and $y_i = 0$ for $i > k$) we conclude that (2.1) is inconsistent.

Particular cases of the duality theorem may be traced back to Gordan [23] and Farkas [16]. The notion of a dual LP problem was introduced by John von Neumann in conversations with George B. Dantzig in October 1947; it appears implicitly in his working paper [36]. Gale, Kuhn and Tucker [19] formulated, and proved, an explicit version of the duality theorem (our "second version"). Our "third version" comes from Kuhn [27]. For a wealth of information on the subject, the reader is referred to Dantzig's book [9].

The duality theorem is a very natural principle, pervading a large area of mathematics. For instance, the necessary and sufficient conditions for solvability of systems of linear equations are just a very special case of the duality theorem. Averaging arguments, counting of pairs in two different ways, and "Lubell's method" illustrated in Section 1, are rudimentary applications of the duality theorem. Like M. Jourdain who, for more than forty years, had been talking prose without any idea of it, we may often be unaware that our arguments rest, in fact, on the duality theorem.

3. Linear Programming as a Methodological Tool

Linear programming problems may come up in various guises. Sometimes their constraints are only implicit in the problem formulation and it may take considerable effort to uncover them. However, once we recognize the linear programming nature of a problem, we gain a valuable guiding principle: the duality theorem. The following case story of a geometrical problem with an underlying LP structure will illustrate the point.

We shall consider the infinite square grid in the ordinary plane; by definition, each cell in this grid has eight neighbors. A coloring of the cells red and blue will be called feasible if

(i) there is at least one blue cell,

(ii) every blue cell has at least six blue neighbors.

Trivially, coloring all the cells blue we obtain a feasible coloring. A nontrivial feasible coloring, constructed by Fejes Tóth [18], is shown in Figure 1. (The cells marked by crosses are red, the unmarked ones are blue.) In this coloring, "four out of every fifteen" cells are red. Introducing the notion of density (as in [17], pp. 161-162), one can make the last statement more precise. To do so, begin with an arbitrary cell; let its Cartesian coordinates be a, b. For every nonnegative integer k, define S_k to be the set of all those $(2k+1)^2$ cells with coordinates i, j that satisfy

$$|i-a| \leq k \;,\; |j-b| \leq k \;.$$

If X is a set of cells then the lower and the upper limit of the sequence

$$\frac{|X \cap S_0|}{|S_0|}, \frac{|X \cap S_1|}{|S_1|}, \ldots, \frac{|X \cap S_i|}{|S_i|}, \ldots$$

do not depend on our choice of a and b. These two limits are called the lower and the upper density of X; if they coincide then their common value is called the density of X. The set of the red cells in Figure 1 has density $4/15$; Fejes Tóth conjectured that the red upper density of a feasible coloring never exceeds $4/15$.

				X	X			X	X					
								X	X			X	X	
	X	X										X	X	
	X	X			X	X								
					X	X			X	X				
									X	X			X	X
		X	X										X	X
		X	X			X	X							
						X	X			X	X			
X										X	X			X
X			X	X										X
			X	X			X	X						
							X	X			X	X		
X	X										X	X		
X	X			X	X									

Figure 1

Familiarizing ourselves with feasible colorings, we find that they cannot contain various clusters of red cells. For instance, if we begin with three red cells in a row (as in Figure 2) then the feasibility constraint (ii) forces us to paint the entire plane red, thereby violating the constraint (i).

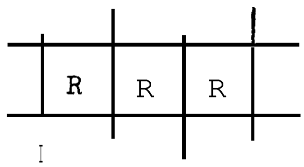

Figure 2

Similarly, we find that no red cell may have more than three red neighbors. In fact, the red cells with exactly three red neighbors come in two by two quadruples flanked by layers of blue cells as in Figure 3. (The cells marked by a questionmark may be red or blue.)

?	B	B	B	B	?
B	B	B	B	B	B
B	B	R	R	B	B
B	B	R	R	B	B
B	B	B	B	B	B
?	B	B	B	B	?

Figure 3

If r (resp. b) is the red upper (resp. blue lower) density of some feasible coloring then trivially

$$r+b = 1 \ .$$

Given a positive ε, we may choose an arbitrarily large $n \times n$ square with at least $(r-\varepsilon)n^2$ red cells. Let N be the number of (unordered) pairs of neighboring cells colored by different colors and coming from our square. Since each red cell has at least five blue neighbors, we have

$$N > 5(r-\varepsilon)n^2 - 4(n+1) \ , \tag{3.1}$$

the negative term discounting the blue cells that fall just outside of our square. Since each blue cell has at most two red neighbors, we have

$$N \leq 2(b+\varepsilon)n^2 \ . \tag{3.2}$$

Since ε may be chosen arbitrarily small and n may be chosen arbitrarily large, we conclude that

$$5r \leq 2b \ . \tag{3.3}$$

Thus we are led to the following LP problem:

$$\left.\begin{array}{l} \text{maximize } r \text{ subject to the constraints} \\ \quad r > 0 \ , \ b \geq 0 \\ \quad r+b = 1 \\ \quad 5r - 2b \leq 0 \end{array}\right\} \tag{3.4}$$

Trivially, the solution to this problem is 2/7 and so every feasible coloring has density at most 2/7 ; unfortunately, 2/7 is just a tiny bit bigger than 4/15. Nevertheless, we may hope that the LP problem (3.4) is, in fact, a poor model of the geometrical problem. To begin with, we may try to prove that there is no feasible coloring with red upper density 2/7 . For this purpose, let us investigate the properties of such a hypothetical coloring.

Since $r = 2/7$ and $b = 5/7$ satisfy (3.3) with the sign of equality, it appears that the bounds (3.1) and (3.2) must be, in some sense, tight. Pursuing this line further we find arbitrarily large squares (100×100 will do nicely) where every red cell has exactly three red neighbors and every blue cell has exactly two red neighbors. Close to the middle of such a square, we shall find the configuration of Figure 3. Next, each of the four cells marked by a questionmark must actually be red (otherwise we would have a blue cell with seven blue neighbors). That is, each of these cells must come from another red quadruple. The blue layers surrounding these quadruples will create blue cells with eight blue neighbors: a contradiction.

The crucial point in our argument was that in the vicinity of each red quadruple, there must be either a red cell with fewer than three red neighbors or a blue cell with more than six blue neighbors. Now that we have established the existence of such defects, we may try to estimate their frequency. For this purpose, we define the order of a red (resp. blue) cell to be the number of its red (resp. blue) neighbors. In a big $n \times n$ square with at least $(r-\varepsilon)n^2$ red cells, let $r_i n^2$ (resp. $b_i n^2$) be the number of red (resp. blue) cells of order i . A careful analysis of the above argument leads to the conclusion that, with only a negligible error,

$$4r_0 + 2r_1 + 2b_7 + 4b_8 \geq r_3 \quad .$$

(The proof of this inequality is not instant; for details, the reader is referred to [7].) In addition, the constraints of (3.4) find their natural counterparts in terms of the new variables. Thus we are led to a new LP problem:

maximize $r_0 + r_1 + r_2 + r_3$ subject to

$$r_0 + r_1 + r_2 + r_3 + b_6 + b_7 + b_8 = 1 ,$$

$$8r_0 + 7r_1 + 6r_2 + 5r_3 - 2b_6 - b_7 = 0 ,$$

$$-4r_0 - 2r_1 + r_3 - 2b_7 - 4b_8 \leq 0 ,$$

$$r_0, r_1, r_2, r_3, b_6, b_7, b_8 \geq 0 .$$

Multiplying the first constraint by four, the second by two, the third by one, and summing up the lot, we arrive at the inequality

$$16r_0 + 16r_1 + 16r_2 + 15r_3 \leq 4 .$$

Hence $r_0 + r_1 + r_2 + r_3 \leq 4/15$ which is the desired result.

4. The Importance of Being Discrete

Reviewing the two examples of Section 1, we find that in the proofs, no use has been made of the fact that our variables x_s were integral. Unfortunately, we cannot expect to get away with that every time we solve a combinatorial problem by LP techniques. Consider, for instance, the problem of finding the largest size of a stable (independent) set of vertices in the graph G = (V,E) of Figure 4. (A set of vertices is called stable if no two of them are joined by an edge.)

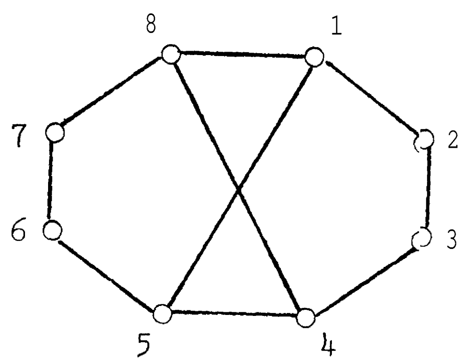

Figure 4

In the straightforward LP formulation of this problem, we have to

$$\text{maximize } \sum_{i=1}^{8} x_i \quad \text{subject to the constraints}$$

$$\left.\begin{array}{ll} x_i \geq 0 & \text{for every vertex } i \text{ ,} \\ x_i + x_j \leq 1 & \text{for every edge } ij \end{array}\right\} \quad (4.1)$$

and

$$x_i = \text{integer for every vertex } i \text{ .} \quad (4.2)$$

Disregarding the integrality constraint (4.2), we may obtain a solution of (4.1) such that

$$\sum_{i=1}^{8} x_i = 4 \quad .$$

However, it is not difficult to see that the size of the largest stable set in G is only three. Conclusion: in the present context, the integrality constraint (4.2) must be taken into account.

In the field of discrete mathematics, we deal with discrete variables; whenever their discreteness is relevant, it must find its way into our arguments. It often does so via the pigeon-hole principle: if mn+1 objects are distributed among n boxes then some box contains at least m+1 objects. Describing this principle in LP terms, we denote the number of objects in box i by x_i ; since the boxes are unlabelled, we may assume that

$$x_1 \geq x_2 \geq \cdot \cdot \cdot \geq x_n \quad .$$

Now, let us

$$\left.\begin{array}{l} \text{minimize } x_1 \text{ subject to the constraints} \\ x_1 - x_2 \geq 0 \\ x_2 - x_3 \geq 0 \\ \cdots \\ x_{n-1} - x_n \geq 0 \\ x_1 + x_2 + \ldots + x_n = mn+1 \end{array}\right\} \quad (4.3)$$

and the constraint

$$x_i = \text{integer} \qquad (1 \le i \le n) \quad .$$

The linear combination of (4.3) with multipliers

$$\frac{n-1}{n}, \frac{n-2}{n}, \ldots, \frac{2}{n}, \frac{1}{n}, \frac{1}{n}$$

reads

$$x_1 \ge m + \frac{1}{n} \quad . \qquad (4.4)$$

At this moment, let the discreteness come into the play: if x_1 is an integer satisfying (4.4) then, in fact, $x_1 \geq m+1$.

Proving, as we have just done, the pigeon-hole principle by LP techniques may be reminiscent of the use of a sledge-hammer to crack the proverbial walnut. We have done so, however, to illustrate a point. The point is that the integrality constraint, together with our linear constraints, may imply inequalities which are not implied by the linear constraints alone. This important idea seems to have appeared for the first time in the work of Dantzig, Fulkerson and Johnson [10]. Later it was developed by Gomory [20],[21],[22] into an algorithm for solving LP problems in integers. Gomory's algorithm provides a systematic way of generating the new "implied" constraints (commonly called cutting planes) until the integrality constraint becomes superfluous. (For an excellent coverage of the ILP techniques, the reader is referred to [30].)

We shall use the idea of implied constraints to formulate a theorem which, in the context of integer LP problems, parallels the duality theorem. To begin with, let

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i \qquad (1 \le i \le m) \qquad (4.5)$$

be a set of inequalities whose solution set is nonempty and bounded. We shall say that (4.5) implies some inequality

$$\sum_{j=1}^{n} c_j x_j \leq d \qquad (4.6)$$

over the integers if every integer solution of (4.5) satisfies (4.6). For instance, the inequalities (4.1) imply

$$\sum_{j=1}^{8} x_j \leq 3 \qquad (4.7)$$

over the integers. When S is a set of linear inequalities, such as (4.5), we define the elementary closure e(S) of S to be the set of all the inequalities

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} \lambda_i a_{ij} \right) x_j \leq d$$

such that

(i) $\lambda_1, \lambda_2, \ldots, \lambda_m$ are nonnegative reals,

(ii) each $\sum \lambda_i a_{ij}$ is an integer,

(iii) d is at least the integer part $\lfloor \sum \lambda_i b_i \rfloor$ of $\sum \lambda_i b_i$.

For instance, the inequality

$$\sum_{j=1}^{5} x_j \leq 2$$

belongs to the elementary closure of (4.1); indeed, the inequality

$$\sum_{j=1}^{5} x_j \leq \frac{5}{2}$$

is a linear combination of (4.1). Clearly, if an inequality belongs to e(S) then S implies this inequality over the integers. However, the converse is not true: for example, it can be shown that the inequality (4.7) does not belong to the elementary closure of (4.1). We shall define $e^0(S) = S$ and, for every positive integer k ,

$$e^k(S) = e(S \cup e^{k-1}(S)) .$$

The set

$$\bigcup_{i=0}^{\infty} e^i(S)$$

will be called the closure of S . Again, it is easy to see that every inequality belonging to the closure of S is, in fact, implied by S over the integers. This time, as asserted by our next theorem, the converse is true. The theorem may be deduced from the finiteness of Gomory's algorithm; a direct proof is given in [5]. (For a thorough analysis of the relationship between Gomory's "fundamental cuts" and our implied constraints, the reader is referred to [32].)

THEOREM. Let S be a set of linear inequalities whose solution is nonempty and bounded. If some linear inequality is implied by S over the integers then this inequality belongs to the closure of S.

For example, if S is the set (4.1) then (4.7) belongs to $e^2(S)$. To see this, consider the inequalites

$$x_1 + x_8 \le 1 ,$$

$$x_2 + x_3 \le 1 ,$$

$$x_4 + x_8 \le 1 ,$$

$$x_5 + x_6 \le 1 ,$$

$$x_6 + x_7 \le 1 ,$$

$$x_7 + x_8 \le 1 ,$$

$$x_1 + x_2 + x_3 + x_4 + x_5 \le 2 .$$

All of them belong to $e^1(S)$; taking their linear combination with multipliers

1/3 , 1/3 , 1/3 , 1/3 , 2/3 , 1/3 , 2/3

we-obtain

$$\sum_{j=1}^{8} x_j < 11 .$$

Since the integer part of 11/3 is 3, we see that the implied constraint (4.7) indeed belongs to $e^2(S)$.

The concept of elementary closure is motivated by the work of Edmonds [12] on the interplay of linear programming and matching theory. When $G = (V,E)$ is a graph, we associate a variable x_j with each edge $j \in E$; a set of edges with a common endpoint is called a star. Let S denote the set of inequalities

$$\begin{aligned} x_j &\geq 0 \quad \text{for every edge } j \in E , \\ \sum_{j \in T} x_j &\leq 1 \quad \text{for every star } T . \end{aligned} \tag{4.8}$$

Clearly, a zero-one vector $(x_j : j \in E)$ satisfies S if and only if it is the characteristic vector of some matching in G . In particular, if m is the size of the largest matching in G then the inequality

$$\sum_{j \in E} x_j \leq m \tag{4.9}$$

is implied over the integers by S . Hence, by the above theorem, (4.9) belongs to the closure of S. In this case, however, a much stronger statement can be made. Indeed, it follows at once from Berge's generalization [3] of Tutte's perfect matching theorem [35] that (4.9) is a linear combination of inequalities from the elementary closure of S . This fact has been pointed out and generalized by Edmonds who proved that the closure of S consists of combinations of $e(S)$.

One may interpret Edmonds' theorem by saying that in matching problems, the integrality constraint is important (it cannot be dropped) but not all that important (unlike (4.2), it may be done away with in just one '*generation" of cutting planes). This interpretation leads to ranking all the integer LP problems according to the "importance" of their integrality constraint. More precisely, when S is a set of linear constraints, we define the rank of S to be the smallest k such that the closure of S consists of combinations of $e^k(S)$.

Let us see how this notion of rank applies to the problem of finding the largest size $a(G)$ of a stable set in a graph $G = (V,E)$; this problem is sometimes called the vertex packing problem. We shall write $V = \{1,2,\dots,n\}$; with each vertex j , we shall associate a variable x_j . Clearly, a zero-one vector $(x_j : j \in V)$ satisfies the constraints

$$\left.\begin{array}{ll} 0 \le x_j \le 1 & \text{for every vertex } j \\ x_i + x_j \le 1 & \text{for every edge } ij \end{array}\right\} \qquad (4.10)$$

if and only if it is the characteristic vector of some stable set in G . The rank of (4.10) is zero if and only if G is bipartite. It is not difficult to find graphs for which (4.10) has arbitrarily high rank. Indeed, if G is a complete graph with n vertices then (4.10) has rank $1 + \lfloor \log_2(n-2) \rfloor$. (The upper bound is not difficult to establish; the lower one follows from Lemma 7.1 of [5].) However, the vertex packing problem is trivial for complete graphs and so (4.10) does not seem to be a well-chosen constraint set. Furthermore, the matching problem for G is the vertex packing problem for the line-graph of G ; however, the constraints (4.8) for G do not reduce into (4.10) for the line graph of G . For these reasons, the stronger set of constraints

$$\left.\begin{array}{ll} 0 \le x_j \le 1 & \text{for every vertex } j \text{ ,} \\ \sum_{j \in C} x_j \le 1 & \text{for every clique } C \end{array}\right\} \qquad (4.11)$$

may be preferred to (4.10). Since (4.10) and (4.11) have the same set of integer solutions, the rank of (4.11) does not exceed that of (4.10); in some cases, it is considerably smaller. For instance, if G is complete then the rank of (4.11) is zero. More generally, (4.11) has rank zero if and only if G is perfect. (This is Theorem 3.1 of [6]. An alternative proof, due to Fulkerson, may be found in [34].) It is not entirely trivial to find graphs with high rank of (4.11) but they do exist.

THEOREM. For arbitrarily large n , there are graphs with n vertices such that the rank of (4.11) is greater than c log n .

For a proof, the reader is referred to [5]. The theorem puts vertex packing problems in a sharp contrast with matching problems: while the latter have rank of most one, there is no upper bound on the rank of the former. In the next section, we shall allude to a theorem which suggests that the vertex packing problems are very hard to solve.

5. Good Algorithms and Good Characterizations

Finally, we turn our attention tomeasuring the difficulty of solving combinatorial problems. In this context, a problem consists of an input together with a "yes or no" question. For example,

$$\left.\begin{array}{ll} \text{Input:} & \text{a graph } G \text{ and an integer } k\ . \\ \text{Question:} & \text{is } \alpha(G) > k\ ? \end{array}\right\} \qquad (5.1)$$

is a problem. Customarily, the size of the input is measured, roughly speaking, by the number of times we must hit the keys of our typewriter in order to describe the input. For instance, a graph G with n vertices may be described by a binary sequence of length at most n^2 ; similarly, the ordinary decimal expansion of a positive integer k has $1+\lfloor \log_{10} k \rfloor$ digits. It has become a common practice to consider a problem solved if there is an efficient algorithm for solving it. In particular, Edmonds [13] pioneered the distinction between "finite" and "better-than-finite" algorithms; he proposed to call an algorithm good if there is a polynomial p such that, given any input of size m , the algorithm terminates within p(m) steps. For instance, Edmonds' algorithm [13] for solving the problem

Input: a graph G and an integer k .
Question: is there a matching of size k ?

is good: indeed, if G has n vertices then the algorithm terminates within $O(n^4)$ steps. On the other hand, no good algorithm for solving "he problem (5.1) is known.

Another important concept, also introduced by Edmonds [11] is that of a good characterization. If we manage to find, by accident or perseverance, a stable set S in G such that $|S| > k$ then we know that the answer to the question in (5.1) is "yes". More importantly, we can use the set S to convince others, in $O(n^2)$ steps, that the answer is "yes". Indeed, there is a good algorithm for solving the problem

Input: a graph $G = (V,E)$, a subset S of V and an integer k .

Question: is S a stable set of size greater than k ?

This fact makes us say that (5.1) has a good characterization. The difference between good algorithms and good characterizations reflects the contrast between the difficulty of finding a solution to a problem and the ease of checking that a proposed solution to a problem is correct.

It may be worthwhile to point out that there is no known good characterization of the problem

$$\left.\begin{array}{ll}\text{Input:} & \text{a graph } G \text{ and an integer } k\,. \\ \text{Question:} & \text{is } \alpha(G) \le k\ ?\end{array}\right\} \qquad (5.2)$$

Indeed, if the answer to this question turns out to be affirmative, we have no easy way of convincing others that this is so. In other words, no efficient way of proving (not to mention finding the proof) that $\alpha(G) \le k$ is known.

How does linear programming fit in this framework? To begin with, no good algorithm for the problem

$$\left.\begin{array}{ll}\text{Input:} & \text{a set } S \text{ of linear inequalities.} \\ \text{Question:} & \text{is } S \text{ solvable?}\end{array}\right\} \qquad (5.3)$$

is known. Indeed, the simplex method (with its standard criteria for column selection), although extremely useful and efficient in practice, takes super-polynomial time on certain artificially constructed examples [26],[37]. Nevertheless, (5.3) does have a good characterization. That is rather obvious: in order to prove that S is solvable, it suffices to exhibit some solution to S . Then it does not take long to verify that the numbers we pulled out of a hat do indeed constitute a solution to S . (To be a little more honest, we should admit that there is a slight catch here. For example, one might proudly present

$$x = 3.14159\ 26535\ 89793\ 23846\ 26433\ 83279\ 50288$$

in order to prove that the inequalities

$$2x - 7 \leq 0$$
$$-8x + 25 < 0$$

are solvable. That would be not only silly, it would also be quite inefficient. Fortunately, whenever S is solvable, at least one of its solutions can be described by a number of digits which does not exceed a certain polynomial in the size of the input.) Less trivially, the, "opposite" of (5.3), that is, the problem

Input: a set S of linear inequalities.
Question: is S unsolvable?

has a good characterization. This fact is just a corollary-to the duality theorem. Indeed, S is unsolvable if and only if it is inconsistent; the inconsistency of S may be proved simply by exhibiting the appropriate multipliers of reasonably small size. 'Let us summarize: (5.3) has a good characterization, its opposite has a good characterization and yet we don't know any good algorithm for solving (5.3). This seems to be a rather rare phenomenon; the only other instance known to the author is the problem

Input: a positive integer n .
Question: is n composite? (5.4)

Trivially, this problem has a good characterization; a good characterization of its opposite (is n a prime?), based on the Lucas-Lehmer heuristic, has been developed by Pratt [31]. Thus we have good characterizations for both (5.4) and its opposite and yet we don't know any good algorithm for solving (5.4). However, there is a reasonable chance that such an-algorithm exists. Quite recently, Miller [29] proved the following: if the Extended Riemann Hypothesis is correct, then there is a good algorithm for testing primality.

Concurrently with finding good algorithms for various combinatorial problems, Edmonds [14],[4],[13] conjectured the nonexistence of good algorithms for other combinatorial problems. (These include the traveling salesman problem, testing graph isomorphism and finding, in a family of triples, the largest subfamily of pairwise disjoint triples.) A few years ago, Cook [8] proved a remarkable theorem whose immediate corollary goes as follows: if there is a good algorithm for (5.1) then there is a good algorithm for every problem that has a good characterization. e conclusion of his corollary is stunningly strong. To appreciate its strength, we may recall that there are problems with a finite characterization but without a finite algorithm. (In other words, there are recursively enumerable sets which are not recursive. The proof may be found in [2], Chapter 4.) By analogy, one may be tempted to conjecture that the same statement holds with "finite" replaced by "good". If this is the case then, by Cook's theorem, there is no good algorithm for (5.1). (At this point, a word of warning may be in order: even though Cook's theorem may be interpreted as evidence that there is no good algorithm for (5.1), it by no means constitutes a proof of the nonexistence of such an algorithm. Edmonds' original conjecture to that effect still remains open. In passing, we may also point out that there is nothing exclusive about (5.1) in Cook's theorem: it may be replaced by many other "difficult" combinatorial problems, such as "Is G hamiltonian?". For an impressive list of such problems, see [1] or [24].)

Another corollary to Cook's theorem states the following: if there is a good characterization for (5.2) then there is a good characterization for every problem whose "opposite" has a good characterization. This conclusion, although not quite as strong as the previous one, may be still found hard to accept; in the rest of this section, we shall speculate about the assumption. In the spirit of integer linear programming, we shall propose a system of inference rules which are strong enough to prove $\alpha(G) \le k$ whenever true. Let $G = (V,E)$ be a graphwith $V = \{1,2,\dots,n\}$. With each vertex i of G, we shall associate a variable x_i, with the graph itself, we shall associate the system of inequalities

$$\left.\begin{array}{ll} 0 \le x_i \le 1 & \text{for every vertex } i \\ x_i + x_j \le 1 & \text{for every edge } ij\ . \end{array}\right\} \qquad (5.5)$$

A system of linear inequalities (in the x_i's) will be called an ILP proof of $\alpha(G) < k$ if

(i) each of these inequalities belongs either to (5.5) or to the elementary closure of previous inequalities in the sequence,

(ii) the last inequality reads $\sum_{i=1}^{n} x_i \leq k$.

For example, if G is as in Figure 4, then

$$x_1 + x_2 \leq 1$$
$$x_2 + x_3 \leq 1$$
$$x_3 + x_4 \leq 1$$
$$x_4 + x_5 \leq 1$$
$$x_1 = x_5 \leq 1$$
$$x_1 + x_2 + x_3 + x_4 + x_5 \leq 2$$
$$x_5 + x_6 \leq 1$$
$$x_6 + x_7 \leq 1$$
$$x_7 + x_8 \leq 1$$
$$x_4 + x_8 \leq 1$$
$$x_1 + x_8 \leq 1$$
$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 \leq 3$$

is an ILP proof of $\alpha(G) \leq 3$. In this case, it can be shown that every ILP proof of $\alpha(G) \leq 3$ takes at least twelve lines. In general, when G is a graph with $a(G) = k$, we shall mean by the complexity $c(G)$ of G the smallest number of inequalities in an ILP proof of $a(G) \leq k$. This notion of complexity is somewhat related to that of rank introduced in the last section.

Indeed, an ILP proof may be arranged into an n-ary tree (rather than a line& sequence) of inequalities, each inequality being in the elementary closure of its immediate descendants. The depth of this tree is at most the rank of (5.5) plus one. Conclusion: if the rank of (5.5) is r then

$$c(G) \leq \sum_{i=0}^{r+1} n^i \quad .$$

This bound may be far from best possible. For instance, if G is complete then $r = 1 + \lfloor \log_2(n-1) \rfloor$ whereas

$$c(G) < \binom{n}{2} + (n-2) . \qquad (5.6)$$

From (5.6), we easily conclude the following: if the rank of (4.11) is s then

$$c(G) < \left(\binom{n}{2} + (n-2)\right) \sum_{i=0}^{s+1} n^i . \qquad (5.7)$$

Unfortunately, s may grow beyond every bound and so 5.7) does not. provide a polynomial upper bound on c(G) .

<u>CONJECTURE</u>. For every polynomial p there is a graph G with n vertices such that c(G) > p(n) .

This conjecture is somewhat related to the conjecture that there is no good characterization for (5.2); the differences between the two go as follows.

1. It is conceivable that the above conjecture is true and yet there is a good characterization for (5.2). (Necessarily, such a characterization would have to use more powerful inference rules than those based on our cutting planes.)

2. It is conceivable that the above conjecture is false and yet the shortest ILP proofs of $a(G) \leq k$ do not provide a good characterization for (5.2). (Necessarily, these shortest ILP proofs would have to involve excessively large coefficients.)

ACKNOWLEDGMENT

I am greatly indebted to Jack Edmonds for many stimulating discussions on the subjects expounded here. In particular, he suggested the notions of rank and complexity to me. D. Avis, J. A. Bondy, G. B. Dantzig, D. E. Knuth, B. Mont-Reynaud and A. F. Veinott, Jr. read preliminary versions of this paper. I am grateful to them for many valuable critical comments.

References

[1] A. V. Aho, J. E. Hopcroft and J. D. Ullman, The Design and Analysis of Computer Algorithms, Addison-Wesley, 1974.

[2] M. A. Arbib, Theories of Abstract Automata, Prentice-Hall, Englewood Cliffs, N. J., 1969.

[3] C. Berge, "Two theorems in graph theory," Proc. Nat. Acad. U.S.A. 43 (1957), 842-844.

[4] R. G. Busacker and T. L. Saaty, Finite Graphs and Networks: An Introduction with Applications, McGraw-Hill, New York, 1965. see Section 6-21.

[5] V. Chvátal, "Edmonds polytopes and a hierarchy of combinatorial problems," Discrete Math. 4(1973), 305-337.

[6] V. Chvátal, "On certain polytopes associated with graphs," J. Combinatorial Theory 18B(1975), 138-154.

[7] V. Chvátal, "On a conjecture of Fejes Tóth," Periodica Math. Hungar., to appear.

[8] S. A. Cook, "The complexity of theorem proving procedures," Proc. 3rd Annual ACM Symposium on Theory of Computing, 151-158.

[9] G. B. Dantzig, Linear Programming and Extensions, Princeton University Press, 1963.

[10] G. B. Dantzig, D. R. Fulkerson and S. M. Johnson, "Solution of a large-scale traveling salesman problem," J. ORSA 2 (1954), 393-410.

[11] J. Edmonds, "Minimum partition of a matroid into independent subsets," J. Res. Nat. Bur. Standards B69(1965), 67-72.

[12] J. Edmonds, "Maximum matching and a polyhedron with 0,1-vertices," J. Res. Nat. Bur. Standards B69(1965), 125-130.

[13] J. Edmonds, "Paths, trees and flowers," Canad. J. Math. 17 (1965), 449-467.

[14] J. Edmonds, "Optimum branchings," J. Res. Nat. Bur. Standards B71 (1967), 233-240.

[15] P. Erdös, Chao Ko and R. Rado, "Intersection theorems for systems of finite sets," Quart. J. Math. Oxford Sec. 12 (1961), 313-320.

[16] J. Farkas, "Über die Theorie der einfachen Ungleichungen," J. Reine Angew. Math. 124 (1902), 1-24.

[17] L. Fejes Tóth, Regular Figures, Pergamon Press, 1964.

[18] L. Fejes Tóth, "Five-neighbor packing of convex discs," Periodica Math. Hungar., to appear.

[19] D. Gale, H. W. Kuhn and A. W. Tucker, "On symmetric games," in Contributions to the Theory of Games (H. W. Kuhn and A. W. Tucker, eds.), Princeton University Press, 1950.

[20] R. E. Gomory, "Outline of an algorithm for integer solutions to linear programs/ Bull. Amer. Math. Soc. 64 (1958), 275-278.

[21] R. E. Gomory, "Solving linear programming problems in integers," in Combinatorial Analysis (R. E. Bellman and M. Hall, Jr., eds.), Proc. Tenth Symp. in Appl. Math.-of the AMS, 1960.

[22] R. E. Gomory, "An algorithm for integer solutions to linear programs/ in Recent Advances in Mathematical Programming (R. L. Graves and P. Wolfe, eds.), McGraw-Hill, 1963.

[23] P. Gordan, "Über die Auflösung linearer Gleichungen mit reelen Coefficienten," Math. Ann. 6 (1873), 23-28.

[24] R. M. Karp, "Reducibility among combinatorial problems," in Complexity of Computer Computations (R. E. Miller and J. W. Thatcher, eds.), Plenum Press, 1972.

[25] G. Katona, "A simple proof of the Erdös - Chao Ko - Rado theorem," J. Combinatorial Theory 13B (1972), 183-184.

[26] V. L. Klee and G. J. Minty, "How good is the simplex algorithm," Boeing Math. Note No. 643, (February 1970).

[27] H. W. Kuhn, "Solvability and consistency for linear equations and inequalities," Amer. Math. Monthly 43 (1956), 217-232.

[28] D. Lubell, "A short proof of Sperner's lemma," J. Combinatorial Theory 1 (1966), 299.

[29] G. R. Miller, "Riemann's hypothesis and tests for primality," Proc. 7th ACM Symp. on Theory of Computing, Albuquerque, N. M., 1975.

[30] G. L. Nemhauser and R. Garfinkel, Integer Programming, Wiley-Interscience, 1972.

[31] V. R. Pratt, "Every prime has a succinct certificate," SIAM J. Computing, to appear.

[32] I. Rosenberg, "On Chvátal's cutting plane in integer linear programming," Zeitschift für Oper. Research, to appear.

[33] E. Sperner, "Ein Satz über Untermengen einer endlichen Menge," Math. Zeitschr. 27 (1928), 544-548.

[34] L. E. Trotter, Jr., "Solution characteristics and algorithms for the vertex-packing problem," Thesis, Cornell University (1973).

[35] W. T. Tutte, "The factorizations of linear graphs," J. London Math. Soc. 22 (1947), 107-111.

[36] J. von Neumann, "On a maximization problem," (manuscript), 1947.

[37] N. Zadeh, "A bad network problem for the simplex method and other minimum cost flow algorithms," Math. Programming 5 (1973), 255-266.