

P. L. IVĂNESCU and S. RUDEANU (București)

## A PSEUDO-BOOLEAN VIEWPOINT ON SYSTEMS OF REPRESENTATIVES

1. Many a problem of graph theory can be formulated and solved as pseudo-Boolean programs, i.e. as mathematical programs with bivalent  $(0, 1)$  variables.

Procedures were developed for:

1. finding the "globally minimizing points" (or simply "minimizing points") of a pseudo-Boolean function (i.e. of a real-valued function with 0-1 variables) with or without constraints;

2. determining the "locally minimizing points" of a pseudo-Boolean function with or without constraints (i.e. those points which fulfil the constraints and have no neighbours fulfilling the constraints and yielding a smaller value to the objective function).

These methods were first elaborated by I. Rosenberg and the authors in [9]; an improved version was given in [10] (see also [11]).

Applications were given to a variety of problems arising in graph theory, among which we mention: the determination of the numbers of internal and external stability, the determination of the chromatic number, the determination of the kernels, of the value of the maximal flow and of the minimal cuts in a network, the determination of the minimal decomposition of partially ordered sets into chains. For these and other applications, see [11].

In the present paper it is shown that several problems concerning systems of representatives may be solved by pseudo-Boolean programming.

2. Let  $E = \{e_1, \dots, e_m\}$  be a set and  $\mathcal{S} = (S_1, \dots, S_n)$  an ordered sequence of subsets  $S_j$  of  $E$ . An ordered sequence  $R = (e_{i_1}, \dots, e_{i_n})$  of elements of  $E$  is called a *system of representatives* for  $\mathcal{S}$  if  $e_{i_j} \in S_j$  ( $j = 1, \dots, n$ ). Then  $e_{i_j}$  is said to represent  $S_j$ .

A system of representatives for  $\mathcal{S}$  is called a *system of distinct representatives* if the elements  $e_{i_k} \in R$  are distinct.

Of course, a necessary condition for the existence of systems of distinct representatives is that  $m \geq n$ .

We recall that a *bipartite graph* is an undirected graph  $G = (N, \rho) = (N', N''; \rho)$  for which the set  $N$  of the nodes is decomposed into two

disjoint sets  $N', N''$  such that  $(N', \varrho)$  and  $(N'', \varrho)$  are totally disconnected subgraphs of  $G$ .

A one-to-one correspondence between two subsets  $A'$  and  $A''$  of  $N'$  and  $N''$ , respectively, is called a *matching* if every two corresponding vertices are linked by an edge. The matching is thus characterized by the set of edges involved in it.

A matching  $M$  will be called *maximal* if there is no other matching  $M'$  properly including all the edges in  $M$ ; the matching will be called *absolutely maximal* if no other matching involves a greater number of edges.

In [12] we have shown that the maximal matchings of a bipartite graph can be determined by pseudo-Boolean methods. Hence, considering the bipartite graph  $G = (E, \mathcal{S}; \epsilon)$  we see that the following theorem holds:

**THEOREM 1.** *An ordered sequence  $R = (e_{i_1}, \dots, e_{i_n})$  is a system of distinct representatives for  $\mathcal{S}$  if and only if the set  $\{(e_{i_1}, S_1), \dots, (e_{i_n}, S_n)\}$  is an absolutely maximal matching of the bipartite graph  $G = (E, \mathcal{S}; \epsilon)$ .*

**COROLLARY 1.** A system of distinct representatives exists if and only if the absolutely maximal matchings of  $G$  involve  $n$  edges each.

**COROLLARY 2.** Pseudo-Boolean procedures may be applied for establishing the existence of the systems of distinct representatives, as well as for their effective detection.

To do this, we associate to each system  $R$  of representatives for  $\mathcal{S}$  a characteristic vector  $(x_{ij})$ , defined as follows:

$$x_{ij} = \begin{cases} 1, & \text{if the element } e_i \text{ is taken as a representative of } S_j, \\ 0, & \text{if } e_i \notin S_j \text{ or if } e_i \in S_j \text{ but } e_i \text{ is not taken as a representative of } S_j. \end{cases}$$

We see that  $x_{ij}$  must satisfy

$$(1) \quad \sum_{i=1}^m x_{ij} = 1 \quad (j = 1, \dots, n).$$

Moreover,  $R$  is a system of distinct representatives if and only if its characteristic vector satisfies (1) and

$$\sum_{j=1}^n x_{ij} \leq 1 \quad (i = 1, \dots, m).$$

However, we prefer to determine the  $x_{ij}$ 's using the method given in [12] for the determination of the absolutely maximal matchings. We obtain thus

**THEOREM 2.** *The family  $\mathcal{S}$  has a system of distinct representatives if and only if the maximum of the pseudo-Boolean function*

$$\sum_{i=1}^m \sum_{j=1}^n x_{ij}$$

under the constraint

$$\sum_{i=1}^m \sum_{h=1}^m \sum_{j=1}^n \sum_{k=1}^n (\delta_i^h + \delta_j^k) x_{ij} x_{hk} = 0$$

is equal to  $n$  (or, equivalently, if and only if the maximum of the unrestricted pseudo-Boolean function

$$(2) \quad \sum_{i=1}^m \sum_{j=1}^n x_{ij} + (p+1) \sum_{i=1}^m \sum_{h=1}^m \sum_{j=1}^n \sum_{k=1}^n (\delta_i^h + \delta_j^k) x_{ij} x_{hk}$$

is equal to  $n$ ), where  $\delta$  is the Kronecker symbol and  $p$  is the number of edges

Example 1. Let  $E = \{1, 2, 3, 4, 5, 6, 7, 8\}$  and suppose  $\mathcal{S}$  is composed of  $S_1 = \{1, 2, 4\}$ ,  $S_2 = \{4\}$ ,  $S_3 = \{1, 2, 5, 8\}$ ,  $S_4 = \{1\}$ ,  $S_5 = \{1, 3, 4, 7, 8\}$ ,  $S_6 = \{4, 6, 7\}$ .

The associated graph  $G = (E, \mathcal{S}; \epsilon)$  is that on Fig. 1.

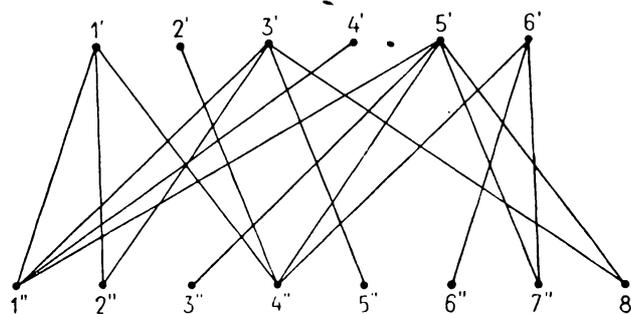


Fig. 1

For the sake of simplicity, let us denote

$$(3) \quad \begin{aligned} x_{11} &= y_1, & x_{12} &= y_2, & x_{14} &= y_3, & x_{24} &= y_4, & x_{31} &= y_5, \\ x_{32} &= y_6, & x_{35} &= y_7, & x_{38} &= y_8, & x_{41} &= y_9, & x_{51} &= y_{10}, \\ x_{53} &= y_{11}, & x_{54} &= y_{12}, & x_{57} &= y_{13}, & x_{58} &= y_{14}, \\ x_{64} &= y_{15}, & x_{66} &= y_{16}, & x_{67} &= y_{17}; \end{aligned}$$

the function (2) becomes

$$(4) \quad \begin{aligned} & \sum_{i=1}^{17} y_i + 18(y_1 y_2 + y_1 y_3 + y_2 y_3 + y_5 y_6 + y_5 y_7 + y_5 y_8 + y_6 y_7 + y_6 y_8 + \\ & + y_7 y_8 + y_{10} y_{11} + y_{10} y_{12} + y_{10} y_{13} + y_{10} y_{14} + y_{11} y_{12} + y_{11} y_{13} + \\ & + y_{11} y_{14} + y_{12} y_{13} + y_{12} y_{14} + y_{13} y_{14} + y_{15} y_{16} + y_{15} y_{17} + y_{16} y_{17} + \\ & + y_1 y_5 + y_1 y_9 + y_1 y_{10} + y_5 y_9 + y_5 y_{10} + y_9 y_{10} + y_2 y_6 + y_3 y_4 + \\ & + y_3 y_{12} + y_3 y_{15} + y_4 y_{12} + y_4 y_{15} + y_{12} y_{15}). \end{aligned}$$

The globally maximizing points of (4) can be determined as in [10]; they are given in Table 1.

TABLE 1

| $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ | $y_9$ | $y_{10}$ | $y_{11}$ | $y_{12}$ | $y_{13}$ | $y_{14}$ | $y_{15}$ | $y_{16}$ | $y_{17}$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|----------|----------|----------|----------|----------|----------|----------|
| 0     | 1     | 0     | 1     | 0     | 0     | 1     | 1     | 1     | 0        | 0        | 0        | 1        | 0        | 0        | 1        | 0        |
| 0     | 1     | 0     | 1     | 0     | 0     | 0     | 1     | 1     | 0        | 0        | 0        | 1        | 0        | 0        | 1        | 0        |
| 0     | 1     | 0     | 1     | 0     | 0     | 1     | 0     | 1     | 0        | 0        | 0        | 0        | 1        | 0        | 1        | 0        |
| 0     | 1     | 0     | 1     | 0     | 0     | 1     | 0     | 1     | 0        | 0        | 0        | 0        | 1        | 0        | 0        | 1        |
| 0     | 1     | 0     | 1     | 0     | 0     | 1     | 0     | 1     | 0        | 1        | 0        | 0        | 0        | 0        | 1        | 0        |
| 0     | 1     | 0     | 1     | 0     | 0     | 1     | 0     | 1     | 0        | 1        | 0        | 0        | 0        | 0        | 0        | 1        |
| 0     | 1     | 0     | 1     | 0     | 0     | 0     | 1     | 1     | 0        | 1        | 0        | 0        | 0        | 0        | 1        | 0        |
| 0     | 1     | 0     | 1     | 0     | 0     | 0     | 1     | 1     | 0        | 1        | 0        | 0        | 0        | 0        | 0        | 1        |

From theorem 2, (3), and Table 1 we see that  $\mathcal{S}$  has eight systems of distinct representatives; they are given in Table 2.

The number of systems of distinct representatives is here *incidentally* equal to the number of elements in  $E$ .

TABLE 2

|       | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_5$ | $S_6$ |
|-------|-------|-------|-------|-------|-------|-------|
| $R_1$ | 2     | 4     | 5     | 1     | 7     | 6     |
| $R_2$ | 2     | 4     | 8     | 1     | 7     | 6     |
| $R_3$ | 2     | 4     | 5     | 1     | 8     | 6     |
| $R_4$ | 2     | 4     | 5     | 1     | 8     | 7     |
| $R_5$ | 2     | 4     | 5     | 1     | 3     | 6     |
| $R_6$ | 2     | 4     | 5     | 1     | 3     | 7     |
| $R_7$ | 2     | 4     | 8     | 1     | 3     | 6     |
| $R_8$ | 2     | 4     | 8     | 1     | 3     | 7     |

Various problems concerning systems of distinct representatives may be translated into the pseudo-Boolean language. For instance, a theorem of Hall [5] states that  $\mathcal{S} = (S_1, \dots, S_n)$  has a system of distinct representatives if and only if for all  $k = 1, 2, \dots, n$  every union of  $k$  sets of  $\mathcal{S}$  contains at least  $k$  distinct elements.

This result yields the following

**THEOREM 3.** *Let  $A = (a_{ij})$  be the adjacency matrix of the graph  $G = (E, \mathcal{S}; \epsilon)$ . Then  $\mathcal{S}$  has*

*a system of distinct representatives if and only if the minimum of the pseudo-Boolean function*

$$m - \sum_{j=1}^n x_j - \sum_{i=1}^m \prod_{j=1}^n (1 - a_{ij} x_j)$$

*is non-negative.*

**Proof.** Let  $\mathcal{S}'$  be a subsystem of  $\mathcal{S}$  and  $(x_1, \dots, x_n)$  its characteristic vector, i.e.

$$x_j = \begin{cases} 1, & \text{if } S_j \in \mathcal{S}', \\ 0, & \text{if } S_j \notin \mathcal{S}'. \end{cases}$$

The element  $e_i \in E$  belongs to the union  $\bigcup_{S_j \in \mathcal{S}'} S_j$  if and only if

$$\prod_{j=1}^n (1 - a_{ij} x_j) = 0.$$

Hence, the number of elements not in  $\bigcup_{S_j \in \mathcal{S}'} S_j$  is equal to

$$\sum_{i=1}^m \prod_{j=1}^n (1 - a_{ij} x_j).$$

On the other hand, the number of sets in  $\mathcal{S}'$  is  $\sum_{j=1}^m x_j$ .

The condition given in Hall's theorem is that the number of elements in  $\bigcup_{S_j \in \mathcal{S}'} S_j$  is not less than the number of sets in  $\mathcal{S}'$ . In other words, this condition requires that the relation

$$m - \sum_{i=1}^m \prod_{j=1}^n (1 - a_{ij} x_j) \geq \sum_{j=1}^m x_j$$

holds for all  $(x_1, \dots, x_n) \in B_2^n$ .

This completes the proof.

**3.** A system  $R$  of representatives for  $\mathcal{S}$  will be called a *system of restricted representatives* with respect to the couples of integers  $d_i, d'_i$ , where  $0 \leq d_i \leq d'_i$  ( $i = 1, 2, \dots, m$ ), if every element  $e_i \in E$  occurs in the system of representatives  $R$  at least  $d_i$  times and at most  $d_i$  times.

The following theorem is obvious:

**THEOREM 4.**  $R$  is a system of restricted representatives for  $\mathcal{S}$  with respect to  $d_i, d'_i$ , if and only if its characteristic vector satisfies

$$(5) \quad \sum_{i=1}^m x_{ij} = 1 \quad (j = 1, 2, \dots, n)$$

and

$$(6) \quad d_i \leq \sum_{j=1}^n x_{ij} \leq d'_i \quad (i = 1, 2, \dots, m).$$

TABLE 3

|        | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|--------|---|---|---|---|---|---|---|---|
| $d_i$  | 2 | 0 | 0 | 2 | 0 | 0 | 1 | 1 |
| $d'_i$ | 4 | 1 | 1 | 3 | 1 | 1 | 1 | 2 |

**Example 2.** Let  $E$  and  $\mathcal{S}$  be defined as in Example 1, and let us seek the systems of restricted representatives for  $\mathcal{S}$  with respect to the numbers  $d_i, d'_i$  given in Table 3.

The system (5) becomes

$$(7) \quad \begin{aligned} y_1 + y_2 + y_3 &= 1, & y_4 &= 1, & y_5 + y_6 + y_7 + y_8 &= 1, \\ y_9 &= 1, & y_{10} + y_{11} + y_{12} + y_{13} + y_{14} &= 1, & y_{15} + y_{16} + y_{17} &= 1, \end{aligned}$$

while the system (6) becomes

$$(8) \quad \begin{aligned} 2 \leq y_1 + y_5 + y_9 + y_{10} &\leq 4, & 0 \leq y_2 + y_6 &\leq 1, & 0 \leq y_{11} &\leq 1, \\ 2 \leq y_3 + y_4 + y_{12} + y_{15} &\leq 3, & 0 \leq y_7 &\leq 1, & 0 \leq y_{16} &\leq 1, \\ 1 \leq y_{13} + y_{17} &\leq 1, & 1 \leq y_8 + y_{14} &\leq 2. \end{aligned}$$

Using the methods given in [10], we find that the system (7), (8) has the solutions which are given in Table 4.

TABLE 4

| $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ | $y_9$ | $y_{10}$ | $y_{11}$ | $y_{12}$ | $y_{13}$ | $y_{14}$ | $y_{15}$ | $y_{16}$ | $y_{17}$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|----------|----------|----------|----------|----------|----------|----------|
| 0     | 0     | 1     | 1     | 0     | 0     | 0     | 1     | 1     | 1        | 0        | 0        | 0        | 0        | 0        | 0        | 1        |
| 1     | 0     | 0     | 1     | 0     | 0     | 0     | 1     | 1     | 0        | 0        | 1        | 0        | 0        | 0        | 0        | 1        |
| 1     | 0     | 0     | 1     | 0     | 0     | 0     | 1     | 1     | 0        | 0        | 0        | 1        | 0        | 1        | 0        | 0        |
| 0     | 0     | 1     | 1     | 1     | 0     | 0     | 0     | 1     | 0        | 0        | 0        | 0        | 1        | 0        | 0        | 1        |

Hence we have found the following systems of restricted representatives with respect to the numbers in Table 3:

$$R_1 = (4, 4, 8, 1, 1, 7), \quad R_2 = (1, 4, 8, 1, 4, 7),$$

$$R_3 = (1, 4, 8, 1, 7, 4), \quad R_4 = (4, 4, 1, 1, 8, 7).$$

4. Ford and Fulkerson [1] have shown that the following condition is necessary and sufficient for the existence of a system of distinct representatives: Every subset  $X$  of the set of indices  $\{1, 2, \dots, n\}$  has no more than

$$(9) \quad \min \left( n - \sum_{i=1}^m d_i + \sum_{i \in I(X)} d'_i, \sum_{i \in I(X)} d'_i \right)$$

elements, where  $I(X) \subset \{1, 2, \dots, m\}$  is the index set of  $\bigcup_{j \in X} S_j$ .

This theorem may be translated into the pseudo-Boolean language as follows:

**THEOREM 5.** *A system of restricted representatives for  $\mathcal{S}$  with respect to  $d_i, d'_i$  exists if and only if the minimum of the pseudo-Boolean function*

$$(10) \quad f(x_0, x_1, x_2, \dots, x_n)$$

$$= x_0 \left[ n + \sum_{i=1}^m d_i \prod_{j=1}^n (1 - a_{ij} x_j) \right] + \bar{x}_0 \sum_{i=1}^m d'_i \left[ 1 - \prod_{j=1}^n (1 - a_{ij} x_j) \right] - \sum_{j=1}^n x_j$$

is non-negative, where  $a_{ij} = 1$  if  $e_i \in S_j$ , and  $a_{ij} = 0$  if  $e_i \notin S_j$ .

**Proof.** Let  $(x_1, x_2, \dots, x_n)$  be the characteristic vector of the set  $X$ . Then, obviously, the following relations hold:

$$(11) \quad n - \sum_{i=1}^m d_i + \sum_{i \in I(X)} d_i = n + \sum_{i \in I(X)} d_i = n + \sum_{i=1}^m d_i \prod_{j=1}^n (1 - a_{ij} x_j),$$

$$(12) \quad \sum_{i \in I(X)} d'_i = \sum_{i=1}^m d'_i \left[ 1 - \prod_{j=1}^n (1 - a_{ij} x_j) \right].$$

Denoting the right-hand sides of (11) and (12) by  $g(x_1, x_2, \dots, x_n)$  and  $h(x_1, x_2, \dots, x_n)$ , respectively, the condition given by Ford and Fulkerson becomes

$$\sum_{j=1}^n x_j \leq \min [g(x_1, x_2, \dots, x_n), h(x_1, x_2, \dots, x_n)]$$

for all  $(x_1, x_2, \dots, x_n) \in B_2^n$ . In other terms,

$$\sum_{j=1}^n x_j \leq x_0 g(x_1, x_2, \dots, x_n) + \bar{x}_0 h(x_1, x_2, \dots, x_n)$$

for all  $(x_0, x_1, x_2, \dots, x_n) \in B_2^{n+1}$ , i.e.

$$\min \left[ x_0 g(x_1, x_2, \dots, x_n) + \bar{x}_0 h(x_1, x_2, \dots, x_n) - \sum_{j=1}^n x_j \right] \geq 0,$$

which coincides with (10).

REMARK. Taking  $d_i = 0$ ,  $\bar{d}_i = 1$  ( $i = 1, 2, \dots, m$ ), Theorem 5 reduces to Theorem 3.

Several other problems concerning systems of representatives (as those on systems of common representatives, marginal elements, matrices of zeros and ones, etc.) may also be handled by means of pseudo-Boolean techniques. As a matter of fact, several proposals were made for using integer linear programming procedures (especially the Gomory algorithm) in order to solve combinatorial problems belonging to the above discussed class. The bibliography concerning systems of representatives is rich enough. For instance, the reader is referred to Ford and Fulkerson [1], [2], Fulkerson and Ryser [3], Hall Jr. [4], Hall [5], Halmos and Vaughan [6], Hoffman and Kuhn [7], [8], Mann and Ryser [13], Mendelsohn and Dulmage [14], Ore [15], and Ryser [16].

#### References

- [1] L. R. Ford, Jr., and D. R. Fulkerson, *Network flows and systems of representatives*, Canad. J. Math. 10 (1958), pp. 78-85.
- [2] —, *Flows in networks*, Princeton University Press, Princeton 1962.
- [3] D. R. Fulkerson and H. J. Ryser, *Widths and heights of (0, 1)-matrices*, Canad. J. Math. 13 (1961), pp. 239-255.
- [4] M. Hall, Jr., *Distinct representatives of subsets*, Bull. Amer. Math. Soc. 54 (1948), pp. 922-926.
- [5] P. Hall, *On representatives of subsets*, J. London Math. Soc. 10 (1935), pp. 26-30.
- [6] P. R. Halmos and H. E. Vaughan, *The marriage problem*, Amer. J. Math. 72 (1950), pp. 214-215.

- [7] A. J. Hoffman and H. W. Kuhn, *On systems of distinct representatives*, in: *Linear inequalities and related systems*, Ann. Math. Studies 38, Princeton University Press, Princeton 1956.
- [8] —, *Systems of distinct representatives and linear programming*, Amer. Math. Monthly 63 (1956), pp. 455-460.
- [9] P. L. Ivănescu, I. Rosenberg and S. Rudeanu, *On the determination of minima of pseudo-Boolean functions* (in Romanian), Studii Cerc. Mat. 14 (1963), pp. 359-364.
- [10] P. L. Ivănescu and S. Rudeanu, *Pseudo-Boolean methods for bivalent programming*, Lecture Notes in Mathematics No. 23, Springer-Verlag, Berlin 1966.
- [11] —, *Boolean methods in operations research and related areas*, Springer-Verlag, Berlin 1968.
- [12] —, *A pseudo-Boolean approach to matching problems in graphs, with applications to assignment and transportation problems*, Actes des Journées d'Études sur la Théorie des Graphes, Rome, July 1966, Dunod, Paris 1967, pp. 161-175.
- [13] H. B. Mann and H. J. Ryser, *Systems of distinct representatives*, Amer. Math. Monthly 60 (1953), pp. 397-401.
- [14] N. S. Mendelsohn and A. L. Dulmage, *Some generalizations of the problem of distinct representatives*, Canad. J. Math. 10 (1958), pp. 230-241.
- [15] O. Ore, *Theory of graphs*, Amer. Math. Soc. Coll. Publ., vol. 38, Providence 1962.
- [16] H. J. Ryser, *Combinatorial mathematics*, The Carus Monographs No. 14, Math. Assoc. of America, 1963.

INSTITUTE OF MATHEMATICS  
ACADEMY OF THE SOCIALIST REPUBLIC OF ROMANIA  
BUCHAREST

*Received on 2. 1. 1967*

---