RAIRO. RECHERCHE OPÉRATIONNELLE

Y. CRAMA
P. L. HAMMER
B. JAUMARD
B. SIMEONE

Product form parametric representation of the solutions to a quadratic boolean equation

RAIRO. Recherche opérationnelle, tome 21, n° 4 (1987), p. 287-305

<http://www.numdam.org/item?id=RO_1987__21_4_287_0>

R.A.I.R.O. Recherche opérationnelle/Operations Research
(vol. 21, n° 4, novembre 1987, p. 287 à 306)

PRODUCT FORM PARAMETRIC REPRESENTATION OF THE SOLUTIONS TO A QUADRATIC BOOLEAN EQUATION (*)

by Y. CRAMA (1), P. L. HAMMER (1),
B. JAUMARD (2) and B. SIMEONE (2) (3)

Abstract. — *A parametric representation of the solutions to a consistent quadratic boolean equation in n variables is obtained. Each variable (or its complement) is expressed as a product of free boolean parameters or their complements. These expressions provide a complete description of the solution set of the equation. An $O(n^3)$ algorithm is proposed to produce such a representation. An application to the maximization of some classes of pseudoboolean functions is discussed.*

Keywords : Quadratic boolean equation, parametric representation, implication graph, transitive closure, complexity.

Résumé. — *On obtient une représentation paramétrique de l'ensemble des solutions d'une équation booléenne quadratique à n variables. Chaque variable (ou son complément) s'exprime comme un produit de paramètres booléens indépendants éventuellement complémentés. L'ensemble de ces expressions décrit complètement l'ensemble des solutions de l'équation. On présente un algorithme en $O(n^3)$ pour obtenir une telle représentation. Une application à la maximisation de différentes classes de fonctions pseudo-booléennes est proposée.*

Mots clés : Équation booléenne quadratique, représentation paramétrique, graphe d'implication, fermeture transitive, complexité.

(*) Received February 1987.
(1) RUTCOR, Rutgers University, New Brunswick, NJ 08903, USA.
(2) RUTCOR and CAIP, Rutgers University, New Brunswick, NJ 08903, USA.
(3) Department of Statistics, University of Rome, « La Sapienza », Italy.

R.A.I.R.O. Recherche opérationnelle/Operations Research, 0399-0559/87/04 287 20/$ 4.00

1. INTRODUCTION

Let $X=\{x_1, x_2, \ldots, x_n\}$ denote a set of $n=|X|$ *boolean variables* and $\bar{X}=\{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n\}$ the set of their *complements*. A *literal* is either a variable or its complement. Literals will be denoted by Greek letters μ, η, ... A *quadratic boolean equation* (E) over X has the form:

$$T_1 \vee T_2 \vee \ldots \vee T_m = 0, \tag{1}$$

where each term T_j can be written as μ or $\mu\eta$, i. e. as either a single literal or a *product* (conjunction) of two literals. Without loss of generality, we shall assume that all terms are products of exactly two literals.

Different graph algorithms have been proposed to check in polynomial time the consistency of a quadratic boolean equation (*see* e. g. Even, Itai and Shamir [1976], Aspvall, Plass and Tarjan [1979], Petreschi and Simeone [1980], Johnson and Padberg [1982], and Simeone [1985]). On the other hand, a quadratic equation can have an exponential number of solutions, and enumerating all of them is in general a prohibitive task. In fact, Valiant [1979] has proved that even determining the number of such solutions is *NP*-complete, and hence probably very difficult.

Nevertheless, as shown in this paper, one can obtain a concise *product form parametric representation* for the set of solutions to an arbitrary quadratic boolean equation. The representation uses no more than n free boolean parameters for an equation in n variables. Each variable (or its complement) is expressed as a product of these parameters or their complements, and these expressions provide a complete description of the solution set of the equation. Furthermore, such representation can be computed in $O(n^3)$ time.

In fact, algebraic methods for determining parametric representations in the case of general boolean equations [i. e. equations of the form (1) with an arbitrary number of literals in each term] have been known for a long time (*see* Löwenheim [1908, 1910] and Rudeanu [1974]). When specialized to quadratic equations, Löwenheim's method produces (in polynomial time) a parametric representation of the solution set, where each variable is associated with some boolean expression of the parameters. The resulting expressions are generally neither in disjunctive nor in conjunctive normal form, and reducing them to such a convenient format can be computationally expensive.

This is to be contrasted with the very simple form of the representation proposed here.

The proposed parametric representation is presented in the next Section, and an algorithm to produce it is described in Section 3. The algorithm relies on the concepts of *implication graph*, defined by Aspvall, Plass and Tarjan [1979], and of *mirror graph*, introduced by Hansen and Jaumard [1985]. Section 4 contains an example. The combinatorial structure of the representation is explored in Section 5, and some applications to 0-1 optimization problems are given in Section 6. Finally, we discuss in Section 7 some properties of the representation.

2. PRODUCT FORM PARAMETRIC REPRESENTATION

For ease of presentation, let us temporarily assume that the quadratic equation (E) given by (1) is *pure*, i. e. all its terms are *positive* (they involve only uncomplemented variables) or *mixed* (they involve exactly one complemented and one uncomplemented variable). This assumption is not very restrictive, since every consistent boolean equation can always be cast into a pure one, by simply renaming some of its variables (*see* Section 3).

Now, for any pair of boolean variables x_k, x_j, the following properties are easily seen to hold:

$$x_k \bar{x}_j = 0 \quad \text{if and only if} \quad x_k \leqq x_j,$$
$$x_k x_j = 0 \quad \text{if and only if} \quad x_k \leqq \bar{x}_j.$$

Therefore, (E) can be rewritten (usually in more than one way) as a system of boolean inequalities of the form:

$$x_k \leqq x_j \qquad (x_j \in D_k) \tag{2}$$
$$x_k \leqq \bar{x}_j \qquad (\bar{x}_j \in D_k) \tag{3}$$

where $D_k \subseteqq X \cup \bar{X}$ $(k = 1, 2, \ldots, n)$. We will also assume that $x_k \notin D_k$, that D_k does not contain both a variable and its complement, and that $x_k \notin D_j$ if $x_j \in D_k$, for $k, j \in \{1, \ldots, n\}$ (these last conditions essentially mean that, in the solution set of the equation, no variable assumes a fixed value, and no two variables can be identified; we will show in Section 3 that there is no loss of generality in assuming so).

The system (2)-(3) is in turn equivalent to the following one:

$$x_k \leqq (\Pi_{x_j \in D_k} x_j)(\Pi_{\bar{x}_j \in D_k} \bar{x}_j) \qquad (k = 1, 2, \ldots, n) \tag{4}$$

and hence also to the system of equations:

$$x_k = x_k (\Pi_{x_j \in D_k} x_j)(\Pi_{\bar{x}_j \in D_k} \bar{x}_j) \qquad (k = 1, 2, \ldots, n). \tag{5}$$

The expression (5) of the equation (E) suggests the following construction. Let $p = (p_1, p_2, \ldots, p_n)$ denote a vector of independent boolean parameters, and define:

$$g_k(p) = g_k(p_1, \ldots, p_n) = p_k (\Pi_{j : x_j \in D_k} p_j)(\Pi_{j : \bar{x}_j \in D_k} \bar{p}_j) \tag{6}$$

for $k = 1, 2, \ldots, n$. Let:

$$Q = \{(g_1(p), \ldots, g_n(p)) : p \in \{0, 1\}^n\}, \tag{7}$$

and denote by S the set of solutions of (E). Then:

PROPOSITION 1: $S \subseteq Q$.

Proof. — If $(x_1, \ldots, x_n) \in S$, then $x_i = g_i(x_1, \ldots, x_n)$ for $i = 1, \ldots, n$. Hence, $(x_1, \ldots, x_n) \in Q$. □

The next Proposition states a necessary and sufficient condition on (E) under which equality holds between S and Q. We first introduce some more notations. We use throughout this paper the graph-theoretic terminology of Berge [1973]. All the graphs we consider are directed. With the formulation (5) of (E), we associate the graph $H = (X \cup \bar{X}, A)$ defined as follows: for all x_k in X and μ in $X \cup \bar{X}$, the arc (x_k, μ) is in A if and only if $\mu \in D_k$. We say that H is *transitive* if the arc (x_k, μ) is in A whenever (x_k, x_j) and (x_j, μ) are in A, for some $j \in \{1, \ldots, n\}$.

PROPOSITION 2: *$S = Q$ if and only if H is transitive.*

Proof. — Assume first that H is transitive. By Proposition 1, we only have to prove that every vector $(g_1(p), \ldots, g_n(p))$ in Q is a solution of (2)-(3).

Let $\bar{x}_j \in D_k$. If $g_k(p) = 1$, then $p_j = 0$, and hence $g_j(p) = 0$. This shows that the inequalities (3) are satisfied by $(g_1(p), \ldots, g_n(p))$.

Let $x_j \in D_k$. If $g_j(p) = 0$, then either (i) $p_j = 0$, or (ii) $p_i = 0$ for some i such that $x_i \in D_j$, or (iii) $p_i = 1$ for some i such that $\bar{x}_i \in D_j$. In case (ii), $x_i \in D_k$ by transitivity of H. Similarly, in case (iii), $\bar{x}_i \in D_k$. Hence, in all cases, $g_k(p) = 0$, and the inequalities (2) are satisfied by $(g_1(p), \ldots, g_n(p))$.

Conversely, assume now that $S = Q$, and that H is not transitive. This means that, for some $x_k, x_j \in X$ and $\mu \in X \cup \bar{X}$, (x_k, x_j) and (x_j, μ) are in A, but (x_k, μ) is not in A. Assume for instance that $\mu \in X$, i. e. $\mu = x_i$ for some

$i \in \{1, \ldots, n\}$ (the proof is similar if $\mu \in \bar{X}$). So, $x_j \in D_k$, and $x_i \in D_j$, but $x_i \notin D_k$. Notice that $i \neq k$, by our assumptions on the system (2)-(3).

Let $p = (p_1, \ldots, p_n)$, where $p_k = 1$, $p_i = 0$, $p_l = 1$ if $x_l \in D_k$, and $p_l = 0$ else (it is easy to check that this is a valid assignment of values to the parameters). Then, $g_k(p) = 1$ and $g_j(p) = 0$. So $(g_1(p), \ldots, g_n(p))$ is not a solution of (2)-(3), and $S \neq Q$. □

So, when H is transitive, the expressions $g_k(p)$ $(k = 1, \ldots, n)$ defined by (6) yield a parametric representation of the solutions of (E). Notice that, even if H is not transitive, (E) can always be transformed into an equivalent equation for which the associated graph is transitive, by adding to it the necessary missing terms. More precisely, if $x_k \leqq x_j$ and $x_j \leqq \mu$ are two inequalities in the system (2)-(3), then inequality $x_k \leqq \mu$ is redundant, and can always be added to the system. Iterating this operation until the resulting graph is transitive is clearly equivalent to computing the *transitive closure* of H. We describe in Section 3 an efficient algorithm to compute a parametric representation of the form (6) for the solutions of an arbitrary quadratic equation.

3. ALGORITHM AND COMPLEXITY

The graph-theoretic algorithm we propose for obtaining a parametric representation of the form (6) for the set of solutions follows the steps outlined in the previous Section and proceeds in five stages:

stage 1: check the equation for consistency;

stage 2: identifications of variables;

stage 3: reduction to a pure equation;

stage 4: fixations of variables;

stage 5: computation of a product form parametric representation.

Stage 1

To represent the quadratic boolean equation (E), we use the concept of implication graph introduced by Aspvall, Plass and Tarjan [1979]. The implication graph $G = (X \cup \bar{X}, U)$ contains two arcs, $(\mu, \bar{\eta})$ and $(\eta, \bar{\mu})$, for each term $\mu\eta$ of (1). If the literal associated with the initial vertex of an arc has value 1 in a solution of the quadratic equation, then the literal associated with its terminal vertex must take the value 1, too. Hence the name of the graph.

We record the following obvious fact for future reference:

PROPOSITION 3: *If all literals are complemented and the orientation of all arcs is reversed in G, then the same graph is obtained.*

Moreover, Aspvall, Plass and Tarjan [1979] proved:

PROPOSITION 4: *The equation (E) is consistent if and only if no vertex $x \in X$ is in the same strongly connected component of G as its complement $\bar{x}$.*

We will assume from now on that (*E*) is consistent, so that the condition in Proposition 4 holds.

Stage 2

The *reduced* implication graph of (*E*) is obtained by shrinking each strongly connected component of *G* into a single vertex. This operation corresponds to identifying all literals in each strongly connected component of *G*: in fact, all such literals must take the same value in every solution to (*E*). Observe that, for every strongly connected component *C* of *G*, there is a "mirror" strongly connected component $\bar{C}$ formed by the complements of the literals in *C*. If we denote by μ and $\bar{\mu}$ the new vertices corresponding to *C* and $\bar{C}$, respectively, then the reduced implication graph is seen to retain the "mirror" property of *G* described in Proposition 3.

Stage 3.

Essentially, this stage consists of an efficient procedure for renaming (renumbering and/or switching) the variables, so that the equation becomes pure.

An important property of the reduced implication graph is the absence of circuits. It is well known that any circuit-free graph $D=(V, A)$ admits a *topological linear ordering*, i.e. a bijection $r: V \to \{1, 2, \ldots, |V|\}$ such that $r(u) < r(v)$ for each arc $(u, v) \in A$. The integer $r(u)$ is called the *rank* of vertex *u*.

If the number of vertices of the reduced implication graph (or, equivalently, the number of strongly connected components of the implication graph) is $2n_R$, we have (*see* Hansen and Jaumard [1985]):

PROPOSITION 5: *There exists a topological linear ordering r of the reduced implication graph such that, for every vertex x:*

$$r(x) + r(\bar{x}) = 2n_R + 1.$$

Given a topological linear ordering of the reduced implication graph as in Proposition 5, we rename $y_1, y_2, \ldots, y_{n_R}$ the vertices of rank less than or equal to n_R, so that y_k is the new vertex of rank k. We let $Y=\{y_1, y_2, \ldots, y_{n_R}\}$ and $\bar{Y}=\{\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_{n_R}\}$. The graph G_M on $Y \cup \bar{Y}$ we obtain in that way is called *mirror graph*. So, G_M is isomorphic to the reduced implication graph.

Define now (E_R) as the "reduced" quadratic boolean equation over Y whose implication graph is G_M. Clearly, (E_R) is a pure equation (*see* Section 2).

Notice that, once a parametric representation of the solutions to (E_R) has been obtained, it is straightforward to derive from it a similar parametric representation for the solutions to (E), using the same set of parameters. Therefore, we may now concentrate on the equation (E_R) only.

Stage 4

Let us associate with (E_R) a graph $H=(Y \cup \bar{Y}, A)$ satisfying the following conditions:

(i) the arc (y_k, y_l) is in H if $y_k \bar{y}_l$ is a mixed term of (E_R) (observe that $k \leqq l$);

(ii) exactly one of the arcs $(y_k, \bar{y}_l)$ or $(y_l, \bar{y}_k)$ is in H if $y_k y_l$ is a positive term of (E_R).

Clearly, H is one of the graphs associated with (E_R) as explained in Section 2, and will be called a *half mirror graph* of (E_R).

We denote by $H^*=(Y \cup \bar{Y}, A^*)$ the *transitive closure* of H: an arc (μ, η) is in A^* if and only if there is a path from μ to η in H. The set of *successors* of the vertex y_k in H^* is denoted by D_k (equivalently, D_k is the set of *descendants* of y_k in H).

PROPOSITION 6: *A variable $y_k \in Y$ is equal to zero in all solutions to (E_R) if and only if either $\bar{y}_k \in D_k$ or $\{y_l, \bar{y}_l\} \subseteq D_k$, for some $l>k$.*

Proof. — The "if" part of the statement is obvious.

For the "only if" part, assume that y_k takes a fixed value in all solutions of (E_R) (this value can only be zero, since (E_R) is pure). It is known (*see* Hansen, Jaumard and Minoux [1986]) that a variable x is fixed in all solutions to a quadratic equation if and only if there exists a path from x to $\bar{x}$, or from $\bar{x}$ to x, in the implication graph of the equation. So, in our case, there exists a path P from y_k to $\bar{y}_k$ in the mirror graph G_M. Let y_i be the last vertex of P which belongs to Y, and $\bar{y}_j$ be the first vertex of P which belongs to $\bar{Y}$. From the definition of G_M, $k \leqq i$ and $k \leqq j$. Also, all the vertices preceding

y_i on P belong to Y, all the vertices following $\bar{y}_j$ on P belong to $\bar{Y}$, and therefore there is an arc from y_i to $\bar{y}_j$ in G_M.

If $i=j=k$, then $\bar{y}_k \in D_k$. If $i=j\neq k$, then $\{y_i, \bar{y}_i\} \subseteq D_k$. So, we may assume that $i\neq j$, and we have $P=(y_k, \ldots, y_i, \bar{y}_j, \ldots, \bar{y}_k)$. By Proposition 3, there must exist in G_M another path $\bar{P}=(y_k, \ldots, y_j, \bar{y}_i, \ldots, \bar{y}_k)$. By symmetry, we may as well assume now that the arc $(y_i, \bar{y}_j)$ is in H. Therefore, the "subpath" $(y_k, \ldots, y_i, \bar{y}_j)$ of P and the "subpath" $(y_k, \ldots, y_j)$ of $\bar{P}$ are paths in H. Hence, $\{y_j, \bar{y}_j\} \subseteq D_k$. □

So, detecting all fixed variables can be done easily once the transitive closure H^* of H is known. We will assume from now on that the vertices associated with such variables have been removed from H^*. For simplicity, we will keep the same notations to describe the new graph we obtain in that way.

Stage 5

Now, let $p=(p_1, p_2, \ldots, p_{n_R})$ denote a vector of independent boolean parameters. Since H^* is transitive, we deduce immediately from Proposition 2:

PROPOSITION 7: *A parametric representation of the solutions to the quadratic equation* (E_R) *is given by:*

$$y_k = g_k(p_k, \ldots, p_{n_R}) = p_k \qquad (\prod_{j : y_j \in D_k} p_j)(\prod_{j : \bar{y}_j \in D_k} \bar{p}_j)$$

for $k=1, 2, \ldots, n_R$.

In summary, the following procedure produces a product-form parametric representation of the solutions to the quadratic boolean equation (E) given by (1).

Step 1 (*Implication graph*). — Construct the implication graph G of (E):

$$G=(X \cup \bar{X}, U).$$

Step 2 (*Strongly connected components*). — Determine the strongly connected components of G.

Step 3 (*Consistency*). — Check if, for some $x \in X$, x and $\bar{x}$ belong to the same strongly connected component of G.

If so, stop: (E) has no solution.

Step 4 (*Mirror graph*). — Construct a mirror graph G_M: shrink each strongly connected component of G into a single vertex; determine a topogical

linear ordering of the new vertices satisfying the condition in Proposition 5; rename the vertices according to their rank.

Step 5 (*Half mirror graph*). – Select a half mirror graph H associated with (E_R), as explained above.

Step 6 (*Transitive closure*). Compute the transitive closure H^* of H.

Step 7 (*Fixation*). – Determine the fixed variables for (E_R), using Proposition 6; delete the corresponding vertices from H^*.

Step 8 [*Parametric representation for* (E_R)]. – Derive a product form parametric representation of the solutions to (E_R), using Proposition 7.

Step 9 [*Parametric representation for* (E)]. – Derive a product form parametric representation of the solutions to (E).

PROPOSITION 8: *The above procedure can be implemented to run in* $O(\max\{m, n_R^3\})$ *time.*

Proof. – Steps 1 and 2 require $O(m)$ operations, using Tarjan's depth-first search algorithm [1972]. In step 3, checking for all $x \in X$ whether both x and $\bar{x}$ belong to a same strongly connected component, takes overall $O(n)$ time. Step 4 can be implemented to run in $O(m)$ time, and step 5 takes $O(m_R)$ time, where m_R is the number of arcs in G_M ($m_R \leqq m$). Hence, this part of the procedure runs in $O(m)$ time. In fact, steps 1 through 4 can be executed simultaneously as shown by Jaumard [1986].

Computing the transitive closure of H, in step 6, can be done in $O(n_R^3)$ operations (*see* e. g. Roy [1959], Warshall [1962] and Mehlhorn [1977]). Determining the fixed variables in step 7 takes $O(n_R^2)$ operations. A parametric representation of the solutions to (E_R) can then be derived in $O(n_R^2)$ time. So, this part of the procedure can be executed in $O(n_R^3)$ time.

Finally, step 9 takes $O(n)$ time.

Thus, the overall time complexity of the procedure is $O(\max\{m, n_R^3\})$. □

4. Example

Consider the quadratic boolean equation:

$$x_1 x_2 \vee x_1 x_3 \vee x_1 \bar{x}_5 \vee x_2 \bar{x}_7 \vee x_3 \bar{x}_4 \vee x_3 \bar{x}_8 \vee x_4 x_5 \times \vee x_4 \bar{x}_7 \vee x_5 x_6 \vee x_6 x_7 \vee x_6 \bar{x}_8 \vee x_7 x_8 = 0 \quad (8)$$

The associated implication graph G is represented in Figure 1.

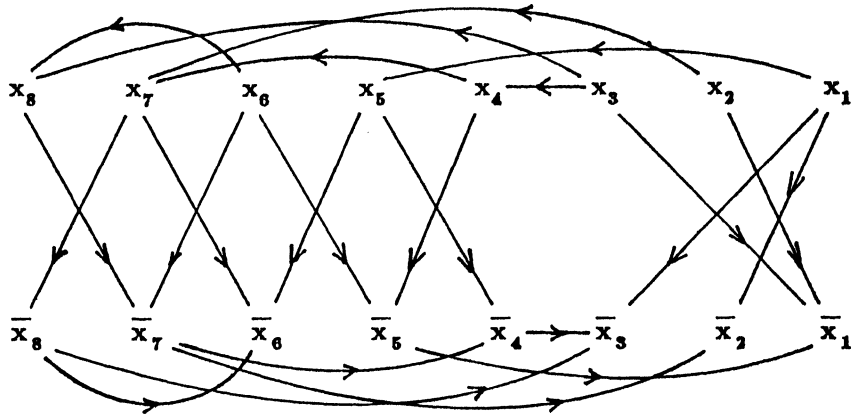

Figure 1. – Implication graph G of the equation (8).

All strongly connected components of G consist of exactly one vertex. Hence, equation (8) is consistent. Moreover, a topological linear ordering of the vertices of G satisfying Proposition 5 is given by:

$$r(x_k)=k, \qquad r(\bar{x}_k)=17-k \qquad (k=1, 2, \ldots, 8)$$

So, we can regard G as being the mirror graph G_M. Then, a half mirror graph H associated with G is shown in Figure 2 and its transitive closure H^* in Figure 3.

At this point, we notice that x_3 must be equal to zero in all solutions of (8), since x_8 and $\bar{x}_8$ are successors of x_3 in H^*.

We are now able to derive a parametric representation of the solutions to (8), under the form:

$$x_1=p_1\bar{p}_2p_5\bar{p}_6$$
$$x_2=p_2p_7\bar{p}_8$$
$$x_3=0$$
$$x_4=p_4\bar{p}_5p_7\bar{p}_8$$
$$x_5=p_5\bar{p}_6$$
$$x_6=p_6\bar{p}_7p_8$$
$$x_7=p_7\bar{p}_8$$
$$x_8=p_8$$

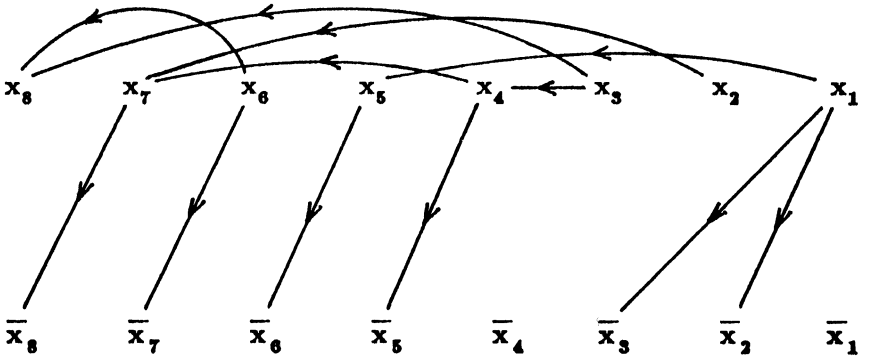

Figure 2. – Half mirror graph H of the equation (8).

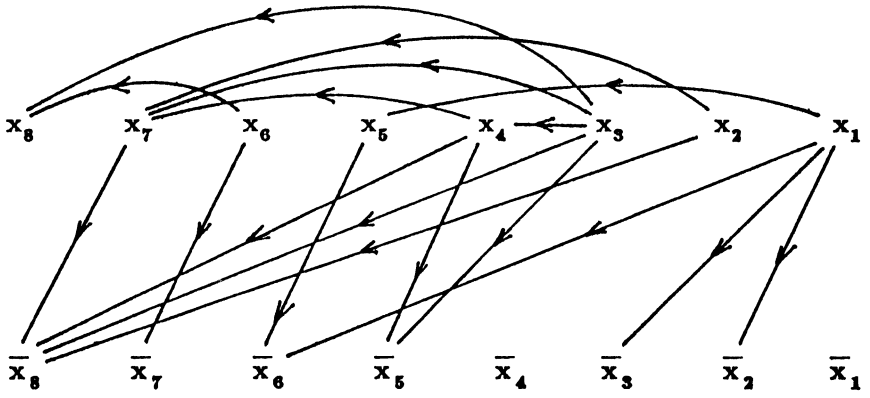

Figure 3. – Transitive closure H^* of H.

Because the mirror graph and the half mirror graph of a given equation are not unique in general, it is usually possible to obtain different parametric representations of the solutions to a given quadratic equation.

For instance, another half mirror graph H' associated with the equation (8) is represented in Figure 4.

The product form representation derived from H′ is the following:

$$x_1 = p_1 \bar{p}_4 p_5$$
$$x_2 = \bar{p}_1 p_2$$
$$x_3 = 0$$
$$x_4 = p_4$$

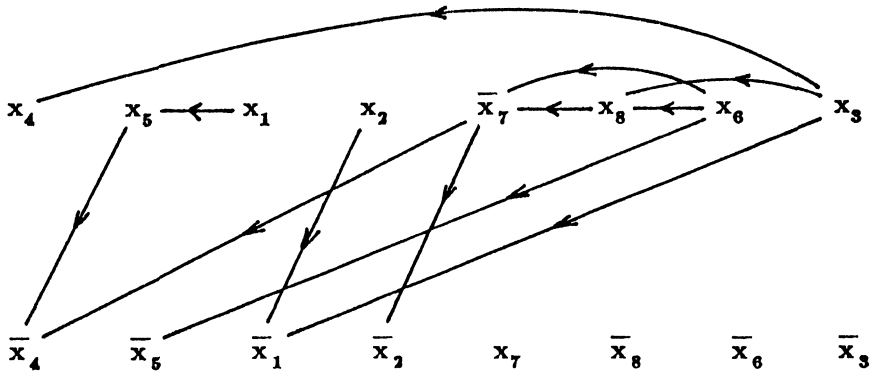

Figure 4. – Half mirror graph H' of the equation (8).

$$x_5 = \bar{p}_4 p_5$$
$$x_6 = \bar{p}_2 \bar{p}_4 \bar{p}_5 p_6 p_7 \bar{p}_8$$
$$\bar{x}_7 = \bar{p}_2 \bar{p}_4 p_7$$
$$x_8 = \bar{p}_2 \bar{p}_4 \bar{p}_7 p_8$$

5. PRODUCT FORM REPRESENTATION: COMBINATORIAL STRUCTURE AND RECOGNITION

Throughout this section, and without loss of generality, we assume that the quadratic boolean equation (E) under consideration satisfies the following conditions: (i) (E) is pure, (ii) there are no fixed variables, (iii) there are no identifications of variables.

In Section 2, it was shown that one can associate with (E) a product form parametric representation:

$$x_k = g_k(p) = p_k (\Pi_{j : x_j \in D_k} p_j)(\Pi_{j : \bar{x}_j \in D_k} \bar{p}_j) \tag{9}$$

Conversely, in this Section we deal with the following *recognition problem: given:* a list of products

$$h_k(p) = \Pi_{p_j \in \Delta_k} p_j \Pi_{\bar{p}_j \in \Delta_k} \bar{p}_j \qquad (k = 1, 2, \ldots, n) \tag{10}$$

question: is there a quadratic equation (E) of the form (5), and a permutation σ of $\{1, 2, \ldots, n\}$ such that

$$h_{\sigma(k)}(p) = g_k(p) \qquad (k = 1, 2, \ldots, n)$$

and $Q = \{(g_1(p), g_2(p), \ldots, g_n(p)) : p \in \{0, 1\}^n\}$ is the set of solutions of (E)?

Observe that as far as the recognition problem is concerned, we may as well assume that the number of parameters is equal to the number of expressions in (10) and that $\Delta_1, \Delta_2, \ldots, \Delta_n$ are all distinct.

The answer to such recognition problem is interesting for different reasons:

(*a*) it will give us some insights on the combinatorial structure of the hypergraph $\{\Delta_1, \Delta_2, \ldots, \Delta_n\}$;

(*b*) it will directly lead to a method for recognizing a new class of pseudo-boolean functions that can be maximized in polynomial time (*cf.* Section 6.2).

Given the expression (10), we introduce the following notations: $P = \{p_1, p_2, \ldots, p_n\}$, $\bar{P} = \{\bar{p}_1, \bar{p}_2, \ldots, \bar{p}_n\}$. For a set $\Delta \subseteq P \cup \bar{P}$, we let $\Delta^+ = \Delta \cap P$ and $\Delta^- = \Delta \cap \bar{P}$.

PROPOSITION 9: *The answer to the recognition problem is affirmative if and only if:*

(C1) $p_i \in \Delta_k \Rightarrow \bar{p}_i \notin \Delta_k$, $(i, k = 1, 2, \ldots, n)$;

(C2) *the hypergraph* $\{\Delta_1^+, \Delta_2^+, \ldots, \Delta_n^+\}$ *has a unique set of distinct representatives (SDR), i.e. there exists a unique permutation* σ *of* $\{1, 2, \ldots, n\}$ *such that* $p_{\sigma(i)} \in \Delta_i^+$ $(i = 1, 2, \ldots, n)$;

(C3) *if* σ *is as in* (2), *then*

$$\Delta_i^+ - \{p_{\sigma(i)}\} = \cup_{j \ \Delta_j \subset \Delta_i} \Delta_j^+ \qquad (i = 1, 2, \ldots, n).$$

Proof. — (*only if*) Assume that (E) is a quadratic equation, that $h_k(p) = g_k(p)$ for all k (this is without loss of generality) and that Q is the set of solutions to (E). Let $\Delta_i = \{p_j : x_j \in D_i\} \cup \{p_i\}$ and let H the graph associated with (E) as in Proposition 2. Then, (C1) holds by our assumptions on (E); (C2) holds since $p_i \in \Delta_i^+$ for all i, and the uniqueness of this (SDR) follows easily from the acyclicity of H; (C3) holds since H is transitive.

(*if*) Assume that (C1), (C2) and (C3) hold and assume for simplicity that $\sigma(i) = i$, $(i = 1, 2, \ldots, n)$. Let us built a graph $H = (X \cup \bar{X}, A)$ in the following way: for $i, j = 1, 2, \ldots, n$,

(i) $(x_i, x_j) \in A$ if $\Delta_j \subset \Delta_i$;

(ii) $(x_i, \bar{x}_j) \in A$ if $\bar{p}_j \in \Delta_i$.

We denote by D_k the set of descendants of vertex x_k in H.

We look at H as a graph associated with some quadratic boolean equation (E). One easily verifies that (E) satisfies all simplifying assumptions made at the beginning of this section.

Also, H is transitive since all arcs in A are defined by inclusion relations. Therefore, by Proposition 2, Q is the set of solutions to (E) and we only have to show that:

$$h_k(p)=g_k(p)=p_k\Pi_{j:x_j\in D_k}p_j\Pi_{j:\bar{x}_j\in D_k}\bar{p}_j \qquad (k=1, 2, \ldots, n) \qquad (11)$$

We prove this by induction on $\delta(k)=|\{j:\Delta_j\subset\Delta_k\}|$. If $\delta(k)=0$, then by condition (C3), $\Delta_k^+-\{p_k\}=\varnothing$, i.e. $\Delta_k^+=\{p_k\}$. Hence:

$$\begin{aligned} g_k(p)&=p_k\Pi_{j:\bar{x}_j\in D_k}\bar{p}_j \\ &=p_k\Pi_{\bar{p}_j\in\Delta_k^-}\bar{p}_j=h_k(p). \end{aligned}$$

Assume now that (11) holds whenever $\delta(k)<i$. In particular, this implies that:

$$p_l\Pi_{j:x_j\in D_l}p_j=\Pi_{p_j\in\Delta_l^+}p_j \qquad (12)$$

for all l such that $\delta(l)<i$.

If $\delta(k)=i$, we get :

$$\begin{aligned} g_k(p)&=p_k(\Pi_{l:x_l\in D_k}p_l)(\Pi_{l:\bar{x}_l\in D_k}\bar{p}_l) \\ &=p_k(\Pi_{l:x_l\in D_k}(p_l\Pi_{j:x_j\in D_l}p_j))(\Pi_{\bar{p}_l\in\Delta_k^-}\bar{p}_l) \end{aligned}$$

(this is true by transitivity of H).

Since $\delta(l)<i$ for all l in D_k, we can use the identity (12) to obtain :

$$\begin{aligned} g_k(p)&=p_k(\Pi_{l:x_l\in D_k}\Pi_{p_j\in\Delta_l^+}p_j)(\Pi_{\bar{p}_l\in\Delta_k^-}\bar{p}_l) \\ &=p_k(\Pi_{l:\Delta_l\subset\Delta_k}\Pi_{p_j\in\Delta_l^+}p_j)(\Pi_{\bar{p}_l\in\Delta_k^-}\bar{p}_l). \end{aligned}$$

So, by (C3) :

$$g_k(p)=(\Pi_{p_l\in\Delta_k^+}p_l)(\Pi_{\bar{p}_l\in\Delta_k^-}\bar{p}_l)=h_k(p),$$

and we are done. □

Observe that Proposition 9 yields an immediate polynomial time algorithm for the solution of the recognition problem.

6. SOME APPLICATIONS

6.1. Degree-two inequalities and maximization of pseudoboolean functions

Johnson, Padberg [1982] and Bourjolly [1983] have investigated "degree-two inequalities" in binary variables, i.e. inequalities of the form $x_i + x_j \geqq 1$, $x_i + x_j \leqq 1$, $x_i \leqq x_j$. Clearly the problem of maximizing a linear function $c(x) = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n$ subject to a set of degree-two inequalities is equivalent to maximizing $c(x)$ subject to a quadratic boolean equation (E).

Let the product form parametric representation of the solutions of (E) be given by:

$$x_i = p_i \Pi_{j\ x_j \in D_i} p_j \Pi_{j\ \bar{x}_j \in D_i} \bar{p}_j \qquad (i \in I)$$

$$\bar{x}_i = p_i \Pi_{j\ x_j \in D_i} p_j \Pi_{j\ \bar{x}_j \in D_i} \bar{p}_j \qquad (i \in \bar{I})$$

where $I \cup \bar{I} = \{1, 2, \ldots, n\}$ and $I \cap \bar{I} = \varnothing$.

Now, replacing each variable in $c(x)$ by its product form parametric representation yields a *pseudoboolean function* $f(p)$ (real-valued function of binary variables) :

$$f(p) = \Sigma_{i \in I} c_i \Pi_{j\ x_j \in D_i} p_j \Pi_{j\ \bar{x}_j \in D_i} (1 - p_j)$$
$$- \Sigma_{i \in \bar{I}} c_i \Pi_{j\ x_j \in D_i} p_j \Pi_{j\ \bar{x}_j \in D_i} (1 - p_j) + \Sigma_{i \in \bar{I}} c_i,$$

and then maximizing $c(x)$ over the solution set of (E) is equivalent to maximizing the pseudoboolean function $f(p)$ over $\{0, 1\}^n$. Notice that this transformation does not increase the number of variables.

6.2. A new class of pseudoboolean functions that can be maximized in polynomial time

An interesting special case of the construction given in Section 6.1 occurs when all the degree-two inequalities are order constraints:

$$x_i \leqq x_j, \qquad (i, j) \in U$$

where $U \subseteqq V \times V$ with $V = \{1, 2, \ldots, n\}$.

Picard [1976] (*see* also Hammer and Simeone [1987]) has shown that maximizing a linear function of binary variables under such constraints is reducible to a minimum cut problem in a network, and hence is solvable in polynomial time.

From Section 6.1 and the discussion in Section 2, it follows that:

$$\text{Max}\{c_1 x_1 + c_2 x_2 + \ldots + c_n x_n : x_i \leqq x_j,\ (i, j) \in U;\ x \in \{0, 1\}^n\} = \text{Max}\{f(p) : p \in \{0, 1\}^n\} \quad (13)$$

where

$$f(p) = \sum_{i=1, n} c_i \Pi_{j \in \Delta_i} p_j, \quad (14)$$

and

$$\Delta_i = \{j : (i, j) \in U\}.$$

Notice that, using the results described by Proposition 9, we can easily recognize in polynomial time those pseudoboolean functions maximization problems arising as described above from the maximization of some linear function subject to a set of order constaints. Indeed, these are exactly the pseudoboolean functions of the form (14), such that $\Delta_1, \ldots, \Delta_n$ satisfy the conditions (C2) and (C3) in Proposition 9 (with $\Delta_i^+ = \Delta_i$, $i = 1, \ldots, n$). This provides a new class of pseudoboolean functions that can be maximized in polynomial time, distinct from other such classes previously introduced e. g. in Barahona [1986], Billionnet and Minoux [1985] or Hansen and Simeone [1986].

6.3. Stable sets in graphs

On the other end of the spectrum, when all degree-two inequalities are of the type $x_i + x_j \leqq 1$, ($\{i, j\} \in U$), the problem stated in Section 6.1 is the well known weighted stability problem for the graph $G = (V, U)$, where $V = \{1, 2, \ldots, n\}$.

Let us define the *pre-neighborhood* of vertex i to be:

$$P_i = \{j : j < i,\ \{i, j\} \in U\}.$$

Specializing the procedure of Section 2, we see that the parametric expressions:

$$x_i = p_i \Pi_{j \in P_i} \bar{p}_j \quad (15)$$

describe precisely the characteristic vectors of the stable sets og G. This construction provides an immediate translation of the weighted stability problem into an unconstrained maximization problem, similar to that used by Ebenegger *et al.* [1984].

7. PROPERTIES OF THE REPRESENTATION

For a consistent quadratic equation (E), the procedure described in Section 2 provides a parametric representation of the solutions to (E), in which every variable or its complement is expressed as a product of free parameters. We discuss now some properties of this product form representation.

1. One might wonder whether such a product form parametric representation exists for boolean equations of order higher than two. The next proposition provides a negative answer to this question.

PROPOSITION 10: *The solution set of a non-quadratic boolean equation cannot, in general, be given a parametric representation in product form.*

Proof. – Suppose for instance that such a representation exists for the equation:

$$x_1 x_2 x_3 \vee \bar{x}_1 \bar{x}_2 \bar{x}_3 = 0. \tag{16}$$

By symmetry, we can assume that the representation is either of the form:

$$x_1 = P_1, \qquad x_2 = P_2, \qquad x_3 = P_3 \tag{17}$$

or of the form

$$\bar{x}_1 = P_1, \qquad x_2 = P_2, \qquad x_3 = P_3 \tag{18}$$

where P_1, P_2 and P_3 denote products of some parameters and their complements.

If the representation is given by (17), then $P_1 P_2 P_3$ must be identically zero. So, we can assume without loss of generality that some parameter p appears uncomplemented in P_1 and complemented in P_2. But then, the solution (1, 1, 0) of (16) is not generated by (17).

So, the representation must be of the form given by (18). Since (0, 0, 1) is a solution of (16), there must exist some literal, say Π_2, that appears in P_2 and not in P_1. Similarly, there must be some literal Π_3 that appears in P_3 and not in P_1. Because $P_1 \bar{P}_2 \bar{P}_3$ is identically zero, $P_1 \bar{\Pi}_2 \bar{\Pi}_3$ must be identically zero too. But, by choice of Π_2 and Π_3, this is only possible if $\Pi_2 = \bar{\Pi}_3$, and then the solution (0, 1, 1) of (16) cannot be generated by (18). □

2. For the sake of simplicity, suppose that no variable takes a constant value, and that no pair of variables take only identical or complementary values, in all solutions of (E). Then:

PROPOSITION 11: *There exists a parametric representation of the solutions to (E) in which every* uncomplemented *variable is expressed as a product of free parameters if and only if (E) is a pure equation.*

Proof. — The "if" part of the statement is an immediate corollary of Proposition 2.

For the "only if" part, assume by contradiction that $\bar{x}_1 \bar{x}_2$ is a term of (E), and that a product form representation of the solutions to (E) is given by:

$$x_i = P_i \qquad (i = 1, 2, \ldots, n),$$

where $P_i (i = 1, 2, \ldots, n)$ is a product of parameters.

From the identity $\bar{P}_1 \bar{P}_2 \equiv 0$, one easily deduces that:

$$P_1 = p \qquad \text{and} \qquad P_2 = \bar{p},$$

for some parameter p. Hence, $x_1 = \bar{x}_2$ in all solutions to (E): contradiction. □

3. Although the particular parametric representation obtained by the procedure of Section 3 depends on the arbitrary choices made in steps 4 and 5 of the algorithm, the number of parameters used is independent of those choices, and is in fact always equal to n_R minus the number of fixed variables for (E_R). It is probably worth pointing out that the solution set of some quadratic equation admits a product form parametric representation using a smaller number of parameters.

For instance, our procedure yields a representation using three parameters for the equation:

$$x_1 x_2 \vee x_1 x_3 \vee x_2 x_3 = 0$$

whereas the following representation uses only two:

$$x_1 = p_1 p_2, \qquad x_2 = p_1 \bar{p}_2, \qquad x_3 = \bar{p}_1 p_2.$$

ACKNOWLEDGMENTS

This research was supported by the Air Force Office of Scientific Research grant AFOSR 0271 and by the National Science Foundation grant ECS 85-03212 to Rutgers University.

REFERENCES

1. B. ASPVALL, M. F. PLASS and R. E. TARJAN, *A Linear-time Algorithm for Testing the Truth of Certain Quantified Formulas*, Information Processing Letters, Vol. 8, 1979, pp. 121-123.
2. F. BARAHONA, *A Solvable Case of Quadratic* 0-1 *Programming*, Discrete Applied Mathematics, Vol. 13, 1986, pp. 23-26.

3. C. BERGE, *Graphs and Hypergraphs*, North-Holland, Amsterdam, 1973.
4. A. BILLIONNET and M. MINOUX, *Maximizing a Supermodular Pseudoboolean Function: a Polynomial Algorithm for Supermodular Cubic Functions*, Discrete Applied Mathematics, Vol. 12, 1985, pp. 1-11.
5. J. M. BOURJOLLY, *An Extension of the König-Egerváry Property to Node-weighted Bidirected Graphs*, CORR 83-39, University of Waterloo, 1983.
6. C. EBENEGGER, P. L. HAMMER and D. DE WERRA, *Pseudo-boolean Functions and Stability of Graphs*, Annals of Discrete Mathematics, Vol. 19, 1984, pp. 83-98.
7. P. L. HAMMER and B. SIMEONE, *Order Relations of variables in* 0-1 *Programming*, Annals of Discrete Mathematics, Vol. 31, 1987, pp. 83-112.
8. P. HANSEN and B. JAUMARD, *Uniquely Solvable Quadratic Boolean Equations*, Discrete Applied Mathematics, Vol. 12, 1985, pp. 147-154.
9. P. HANSEN, B. JAUMARD and M. MINOUX, *A Linear Expected-time Algorithm for Deriving all Logical Conclusions Implied by a Set of Boolean Inequalities*, Mathematical Programming, Vol. 34, 1986, pp. 223-231.
10. P. HANSEN and B. SIMEONE, *Unimodular Functions*, Discrete Applied Mathematics, Vol. 14, 1986, pp. 269-281.
11. S. EVEN, A. ITAI and A. SHAMIR, *On the Complexity of Timetable and Multicommodity Flow Problems*, SIAM Journal on Computing, Vol. 5, 1976, pp. 691-703.
12. B. JAUMARD, *Extraction et utilisation de relations booléennes pour la résolution des programmes linéaires en variables* 0-1, Thèse de Doctorat, École Nationale Supérieure des Télécommunications, Paris, 1986.
13. E. L. JOHNSON and M. W. PADBERG, *Degree two Inequalities, Clique Facets, and Biperfect Graphs*, Annals of Discrete Mathematics, Vol. 16, 1982, pp. 169-187.
14. L. LÖWENHEIM, *Uber das Auflösungsproblem im logischen Klassenkalkul*, Sitzungsberichte der Berliner Mathematische Gesellschaft, Vol. 7, 1908, pp. 89-94.
15. L. LÖWENHEIM, *Uber die Auflösung von Gleichungen im logischen Gebietkalkul*, Mathematische Annalen, Vol. 68, 1910, pp. 169-207.
16. K. MEHLHORN, *Data structures and algorithms 2: Graph algorithms and NP-completeness*, Springer-Verlag, Berlin, 1984.
17. R. PETRESCHI and B. SIMEONE, *A Switching Algorithm for the Solution of Quadratic Boolean Equations*, Information Processing Letters, Vol. 2, 1980, pp. 193-198.
18. J. C. PICARD, *Maximal Closure of a Graph and Applications to Combinatorial Problems*, Management Science, Vol. 22, 1976, pp. 1268-1272.
19. B. ROY, *Transitivité et connexité*, Compte-rendus de l'Académie des Sciences de Paris, Vol. 249, 1959, pp. 216-218.
20. S. RUDEANU, *Boolean Functions and Equations*, North-Holland, Amsterdam, 1974.
21. B. SIMEONE, *Consistency of Quadratic Boolean Equations and the König-Egerváry Property for Graphs*, Annals of Discrete Mathematics, Vol. 25, 1985, pp. 281-290.
22. R. E. TARJAN, *Depth-first Search and Linear Graph Algorithms*, SIAM Journal on Computing, Vol. 1, 1972, pp. 146-160.
23. L. G. VALIANT, *The Complexity of Enumeration and Reliability Problems*, SIAM Journal on Computing, Vol. 8, 1979, pp. 410-421.
24. S. WARSHALL, *A Theorem on Boolean Matrices, Journal of the Association for Computing Machinery*, Vol. 9, 1962, pp. 11-12.