Linear Equations and Interpolation in Boolean Algebra

Robert A. Melter
Department of Mathematics
Southampton College of Long Island University
Southampton, New York 11963

and

Sergiu Rudeanu
Faculty of Mathematics
University of Bucharest
Bucharest, Romania

Submitted by J. J. Seidel

ABSTRACT

The authors generalize the classical interpolation formula for Boolean functions of n variables. A characterization of all interpolating systems with 2^n elements is obtained. The methods of proof used are intimately related to the solution of linear Boolean equations.

I. INTRODUCTION

The term "linear Boolean equation" has been given two different meanings.

The first one coincides with the classical concept of linear equation as applied to the ring structure of a Boolean algebra. In particular, systems of linear Boolean-ring equations have been studied by Parker and Bernstein [2]; see also [3, Chapter 6, §3].

The second meaning of linearity in Boolean algebra is suggested by a simple analogy and is due to Lowenheim [1]. Let $(B,V,\cdot,',0,1)$ be an arbitrary Boolean algebra, where $V,\cdot,',0$ and 1 denote disjunction, conjunction, negation, and the first and last elements, respectively; iterated disjunction and conjunction will be denoted by $\bigvee$ and $\prod$. Lowenheim studied

52 Vanderbilt Ave., New York, NY 10017 0024-3795/84/$3.00

systems of equations of the form

$$\bigvee_{j=1}^{m} a_{ij}x_j = b_i \qquad (i=1,\dots,m),$$

for which he established the essentials of a theory similar to linear algebra. One of us has followed this line of research [3–5], and in the present paper we complete the investigation of linear systems with unique solutions.

We first recall the Muller-Lowenheim verification theorem, which states that if $g_i, h_i\colon B^p \to B$ $(i=1,\dots,m)$ are Boolean functions, then the system of relations

$$g_i(x_1,\dots,x_p)\,\rho_i\,h_i(x_1,\dots,x_p) \qquad (i=1,\dots,m)$$

(where each ρ_i is $=$ or $\leqslant$) holds identically in B if and only if it is valid in the subalgebra $\{0,1\}$ of B (see e.g. [3, Theorem 2.13]).

Let B be a Boolean algebra and n a positive integer, both arbitrary but fixed in the sequel.

In this section we obtain a necessary and sufficient condition for a system of linear Boolean equations

$$\bigvee_{j=1}^{n} a_{ij}x_j = b_i \qquad (i=1,\dots,m) \tag{1}$$

to have a unique solution for a given $(b_1,\dots,b_m)\in B^m$ (Theorem 1) and for each $(b_1,\dots,b_m)\in B^m$ (Theorem 2).

In the Section II we apply these results to obtain a generalization of the interpolation formula for Boolean functions $f\colon B^n \to B$, where by a Boolean function we mean one built up from variables and constants by superpositions of disjunction, conjunction, and negation; cf. [3].

A vector $(x_1,\dots,x_n)\in B^n$ is called: *orthogonal* if $x_ix_j = 0$ for $i \neq j$, *normal* if $\bigvee_{i=1}^{n} x_i = 1$, and *orthonormal* if it is both orthogonal and normal. A Boolean matrix is said to be: *row* or *column orthogonal* (*normal*, *orthonormal*) according as every row or every column is orthogonal (normal, orthonormal), and *orthogonal* (*normal*, *orthonormal*) if it is both row and column orthogonal (normal, orthonormal). It is known (see e.g. Theorem 7.5 in [3]) that if a square matrix is row normal and column orthogonal or row orthogonal and column normal, then that matrix is orthonormal.

We recall that if $c_1,\ldots,c_n$, $x_1,\ldots,x_n \in B$ and $(x_1,\ldots,x_n)$ is orthonormal, then

$$\left(\bigvee_{j=1}^{n} c_j x_j\right)' = \bigvee_{j=1}^{n} c_j' x_j.$$

A system of linear Boolean equations

$$\bigvee_{j=1}^{n} a_{ij} x_j = b_i \qquad (i=1,\ldots,m) \tag{1}$$

is consistent if and only if

$$b_i \leqslant \bigvee_{j=1}^{n} a_{ij} \prod_{\substack{h=1\\ h\neq i}}^{m} (a_{hj}' \vee b_h) \qquad (i=1,\ldots,m) \tag{2}$$

([1]; see also [3, Theorem 7.6]), in which case the set of solutions is given by

$$\bigvee_{i=1}^{m} b_i \left(\prod_{k=1}^{j-1} a_{ik}'\right) \prod_{k=j+1}^{n} (a_{ik}' \vee x_k') \leqslant x_j \leqslant \prod_{i=1}^{m} (a_{ij}' \vee b_i) \quad (j=n,n-1,\ldots,1), \tag{3}$$

as was proved in [4]. More precisely, every sequence $x_n, x_{n-1},\ldots,x_1$ constructed recursively so as to fulfil (3) is a solution of (1), and conversely, every solution of (1) can be obtained in this way.

Now let $\|a_{ij}\|$ be an $m \times n$ Boolean matrix, and regard (1) as a system of equations in which $b_1,\ldots,b_m$ are parameters. If the system (1) is consistent for every $(b_1,\ldots,b_m) \in B^m$, then $m \leqslant n$ [5, Proposition 1]. For $m = n$ the following stronger result holds [5, Theorem 1]:

(i) The system (1) is consistent for every $(b_1,\ldots,b_m) \in B^m$ if and only if the matrix $\|a_{ij}\|$ is orthonormal.

(ii) When this is the case, for each $(b_1,\ldots,b_m) \in B^m$ the unique solution of (1) is

$$x_j = \prod_{i=1}^{m} (a_{ij}' \vee b_i) \qquad (j=1,\ldots,n). \tag{4}$$

Now we determine systems (1) with unique solutions in the case of arbitrary dimensions m, n.

THEOREM 1. *Let $\|a_{ij}\|$ be an $n \times n$ matrix and (b_i) an m-vector over B.*

(i) *The system* (1) *has a unique solution if and only if*

$$\bigvee_{i=1}^{m} b_i\left(\prod_{k=1}^{j-1} a'_{ik}\right)\prod_{k=j+1}^{n}\left(a'_{ik}\vee\bigvee_{h=1}^{m} a_{hk}b'_h\right)=\prod_{i=1}^{m}(a'_{ij}\vee b_i)\quad (j=1,\ldots,n). \tag{5}$$

(ii) *When this is the case, the unique solution is* (4).

Proof. (ii): Suppose $(x_1,\ldots,x_n)$ is the unique solution of the system (1) but $x_j \neq \prod_{i=1}^{m}(a'_{ij}\vee b_i)$ for some $j \in \{1,\ldots,n\}$. As $x_n,\ldots,x_1$ fulfil (3), it follows that $x_j < \prod_{i=1}^{m}(a'_{ij}\vee b_i) = y_j$; hence $x_n,\ldots,x_{j+1}, y_j$ still verify the corresponding inequalities (3), and we can determine successively $y_{j-1},\ldots,y_1$ so that $x_n,\ldots,x_{j+1}$ fulfil all of the inequalities (3), obtaining thus a solution $(y_1,\ldots,y_j, x_{j+1},\ldots,x_n)$ of (1) distinct from $(x_1,\ldots,x_n)$. This contradiction proves that the relations (4) hold.

(i): Note first that the conditions (5) are obtained from (3) by performing the substitution (4) and by replacing each sign $\leqslant$ with $=$. Now suppose that the system (1) has a unique solution. Then the solution is necessarily (4) by (ii), and it fulfils the corresponding conditions (3). Thus a_{ij} and b_i fulfil the conditions obtained from (3) by the substitution (4); moreover, each of the inequalities obtained in this way is in fact an equality: otherwise we could obtain a solution of (1) distinct from (4) by a construction similar to the one used in the above proof of (ii). We have thus established the validity of (5).

Conversely, suppose the conditions (5) hold. This shows that the vector given by (4) fulfils (3); hence it is a solution of (1). Now let $(x_1,\ldots,x_n)$ be an arbitrary solution of (1). Then

$$\bigvee_{i=1}^{m} b_i \prod_{k=1}^{n-1} a'_{ik} \leqslant x_n \leqslant \prod_{i=1}^{m}(z'_{in}\vee b_i)$$

by the last inequality (3), while

$$\bigvee_{i=1}^{m} b_i \prod_{k=1}^{n-1} a'_{ik} = \prod_{i=1}^{m}(a'_{in}\vee b_i)$$

by the last equality (5); hence $x_n=\prod_{i=1}^{m}(a'_{in}\vee b_i)$. Now suppose that the relations (4) hold for $n, n-1,\dots,j+1$. It follows that the left side of the jth inequality (3) coincides with the left side of the jth equality (5), so that, reasoning as above, we deduce easily that $x_j=\prod_{i=1}^{m}(a'_{ij}\vee b_i)$. We have thus proved that the conditions (4) hold for all j, i.e., the unique solution is (4). ■

Theorem 2. *The following conditions are equivalent for an $m\times n$ matrix $\|a_{ij}\|$ over B:*

(i) *The system* (1) *has a unique solution for each* $(b_1,\dots,b_m)\in B$.

(ii) *For every* $M\subseteq\{1,\dots,m\}$,

$$\bigvee_{i\in\overline{M}}\left(\prod_{k=1}^{j-1}a'_{ik}\right)\prod_{k=j+1}^{n}\left(a'_{ik}\vee\bigvee_{h\in M}a_{hk}\right)=\prod_{i\in M}a'_{ij}\qquad(j=1,\dots,n),\quad(6)$$

where $\overline{M}=\{1,\dots,m\}\setminus M$.

Proof. It follows from Theorem 1 that condition (i) is equivalent to the fact that the relations (5) hold for every $(b_1,\dots,b_m)\in B^m$; in view of the Müller-Löwenheim verification theorem, the latter property is equivalent to the fact that the relations (5) hold for every $(b_1,\dots,b_m)\in\{0,1\}^m$, and this is precisely condition (ii). ■

Corollary 1. *Let $\|a_{ij}\|$ be an $m\times n$ matrix over B such that the system* (1) *has a unique solution for every* $(b_1,\dots,b_m)\in B^m$. *Then*:

(i) *The matrix is column normal.*

(ii) *If, moreover, the matrix is row orthogonal, then $m=n$ and the matrix is orthonormal.*

Proof. (i): Taking $M=\{1,\dots,n\}$ in the conditions (6), we obtain

$$0=\prod_{j=1}^{m}a'_{ij}\qquad(j=1,\dots,n).$$

(ii): Column normality and row orthogonality imply $m\geqslant n$ by Lemma 3 of [5] applied to the transpose of $\|a_{ij}\|$, while $m\leqslant n$ follows by Proposition 1 from [5]. Thus $m=n$, and hence the matrix is orthonormal by a known theorem (see e.g. [3, Theorem 7.5]).

COROLLARY 2. *A square Boolean matrix $\|a_{ij}\|$ is orthonormal if and only if it fulfils the conditions* (6) *with $m = n$.*

Proof. Apply Theorem 1 from [5] and Theorem 2. ■

II. INTERPOLATING SYSTEMS

In the sequel we apply the foregoing results to the study of Boolean functions, much in the same line as in [5]. In that paper a *linear generator* (*linear base* or *system of generalized minterms*) for Boolean functions of n variables was defined as a system $(f_i)_{i=1,\ldots,m}$ of Boolean functions $f_i\colon B^n \to B$ $(i=1,\ldots,m)$ such that every Boolean function $f\colon B^n \to B$ can be written in the form

$$f = c_1 f_1 \vee \cdots \vee c_m f_m, \tag{7}$$

where $c_1,\ldots,c_m$ are constants from B (uniquely) determined by f. Every linear generator has at least 2^n functions and for $m = 2^n$ linear generators coincide with linear bases and have been described by several equivalent conditions in Theorem 2 of [5].

One of these linear bases is, of course, the classical system of minterms

$$X^A = x_1^{\alpha_1}\cdots x_n^{\alpha_n}, \qquad A = (\alpha_1,\ldots,\alpha_n) \in \{0,1\}^n, \tag{8}$$

where $x^1 = x$ and $x^0 = x'$. The formula (8) establishes, in fact, a bijection between the 2^n minterms X^A and the vectors $A \in \{0,1\}^n$. As is well known, in this case the coefficients c_i in (7) are determined by

$$c_A = f(A), \qquad A \in \{0,1\}^n; \tag{9}$$

in other words, for every possible choice of the vector $(c_A)_{A\in\{0,1\}^n}$ there is a unique Boolean function f satisfying (9).

This property suggests the following

DEFINITION. A set

$$\Xi = \{\Xi_i\}_{i=1,\ldots,m} \tag{10}$$

of m vectors

$$\Xi_i = (\xi_{i1},\ldots,\xi_{in}) \qquad (i=1,\ldots,m) \tag{11}$$

from B^n will be called an *interpolating system* of points for Boolean functions of n variables, provided every function $g: \Xi \to B$ can be uniquely extended to a Boolean function $f: B^n \to B$.

In the sequel we generalize (8) by establishing a bijection between generalized systems of 2^n minterms and interpolating systems of points (Theorem 3). Then we obtain the corresponding generalization of (9). This yields, via (7), an explicit construction of the Boolean function having prescribed values on a given interpolating system of points (the Corollary of Theorem 3).

Recall first [5, Theorem 2] that a system $(f_i)_{i=1,\dots,2^n}$ is a linear generator or, equivalently, a linear base, if and only if the $2^n \times 2^n$ Boolean matrix

$$\|f_i(A)\|_{i=1,\dots,2^n; A \in \{0,1\}^n} \tag{12}$$

is orthonormal. This holds if and only if the functions f_i are of the form

$$f_i(x_1,\dots,x_n) = \prod_{j=1}^{n} (x_j + \xi_{ij}) \qquad (i=1,\dots,2^n), \tag{13}$$

where $x + y = xy' \vee x'y$ and the vectors (11) are such that the vectors

$$(\Xi_1^A,\dots,\Xi_{2^n}^A) \qquad (A \in \{0,1\}^n) \tag{14}$$

are orthogonal (or, equivalently, normal; or, equivalently, orthonormal).

Lemma. *The following conditions are equivalent for a set* (10) *of vectors from* B^n:

(i) $\{\Xi_i\}_{i=1,\dots,m}$ *is an interpolating system;*
(ii) $m = 2^n$ *and the vectors* (14) *are orthogonal;*
(iii) $m = 2^n$ *and the vectors* (14) *are normal;*
(iv) $m = 2^n$ *and the vectors* (14) *are orthonormal.*

Proof. The condition for the set (10) to be an interpolating system is that for each $(b_1,\dots,b_m) \in B^m$, the system of equations

$$f(\Xi_i) = b_i \qquad (i=1,\dots,m)$$

or, equivalently,

$$\bigvee_{A \in \{0,1\}^n} f(A)\Xi_i^A = b_i \qquad (i=1,\dots,m) \tag{15}$$

have a unique solution $\{f(A)\}_{A\in\{0,1\}^n}$. Note that (15) is a system of the form (1), with the row orthonormal $m\times 2^n$ matrix of coefficients

$$\|\Xi_i^A\|_{i=1,\ldots,m;A\in\{0,1\}^n}. \tag{16}$$

Now suppose (i). Then Corollary 1 of Theorem 2 shows that $m=2^n$. It follows by Theorem 1 of [5] that the coefficient matrix (16) of the system (15) is orthonormal. Further define 2^n functions f_i by (13). Then

$$f_i(A')=\Xi_i^A \qquad (i=1,\ldots,2^n;\quad A\in\{0,1\}^n); \tag{17}$$

hence the matrix (12) is obtained from the orthonormal matrix (16), where $m=2^n$, by a column permutation; therefore the matrix (12) itself is orthonormal. Thus conditions (ii)–(iv) are fulfilled, by Theorem 2 of [5].

Conversely, suppose the vectors Ξ_i fulfil one of the conditions (ii)–(iv). Again define the functions f_i by (13). Then Theorem 2 of [5] shows that all the conditions (ii)–(iv) are fulfilled and, besides, the matrix (12) is orthonormal. It follcws by Theorem 1 of [5] that the system (15) has a unique solution for each $(b_1,\ldots,b_{2^n})\in B^{2^n}$; therefore (i) holds. ■

THEOREM 3. *There is a bijection between interpolating systems of points and systems of* 2^n *generalized minterms, given by*

$$f_i(X)=\prod_{j=1}^{n}(x_j+\xi_{ij}) \qquad (i=1,\ldots,2^n) \tag{13}$$

and

$$\xi_{ij}=\bigvee\left(f_i(A')|A\in\{0,1\}^n\ \&\ \alpha_j=1\right) \qquad (i,\ j=1,\ldots,2^n). \tag{18}$$

Proof. If (10) is an interpolating system, then $m=2^n$ and the vectors (14) are orthonormal by the Lemma; hence the functions f_i defined by (13) form a system of generalized minterms by Theorem 2 of [5]. Thus (13) defines a mapping from interpolating systems to linear bases having 2^n elements. To complete the proof, we show that every 2^n-element linear base is of the form (13), where the vectors Ξ_i $(i=1,\ldots,2^n)$ form an interpolating system and are uniquely determined by (18).

But every system of 2^n generalized minterms is actually of the form (13), with orthonormal vectors (14), again by Theorem 2 of [5]. It follows by the

Lemma that $\{\Xi_i\}_{i=1,\ldots,2^n}$ is an interpolating system. On the other hand, (13) implies (17); hence

$$\bigvee\left(f_i(A')|A \in \{0,1\}^n \; \& \; \alpha_j = 1\right) = \bigvee\left(\Xi_i^A|A \in \{0,1\}^n \; \& \; \alpha_j = 1\right)$$
$$= \xi_{ij} \bigvee\left(C|C \in \{0,1\}^{n-1}\right) = \xi_{ij}.$$

Thus if $(\Xi_i)_{i=1,\ldots,m}$ is an interpolating system, Corollary 1 of Theorem 2 shows that $m = 2^n$. Now Theorem 1 of [5] shows that the matrix (16) of coefficients of the system (15) is orthonormal. Defining the functions f_i $(i = 1,\ldots,2^n)$ by the formulas (10), it follows from (13) that the matrix $\|f_i(A)\|$ is also orthonormal; therefore conditions (i)–(iii) in the statement of Theorem 3 are equivalent and are fulfilled by the vectors Ξ_i $(i = 1,\ldots,2^n)$, in view of Theorem 2 of [5].

Conversely, suppose the vectors Ξ_i $(i = 1,\ldots,2^n)$ fulfil one of the conditions (i)–(iii). As the matrix (16) is row orthonormal, Theorem 7.5 of [3] implies that Ξ is orthonormal; hence Theorem 1 of [5] shows that the system (15) has a unique solution for each $(b_1,\ldots,b_{2^n}) \in B^{2^n}$. Therefore $(\Xi_i)_{i=1,\ldots,2^n}$ is an interpolating system. ■

Corollary 2. *Let $(\Xi_i)_{i=1,\ldots,2^n}$ be an interpolating system. For every Boolean function f and every point $X \in B^n$,*

$$f(X) = \bigvee_{i=1}^{2^n} f(\Xi_i) \bigvee_{A \in \{0,1\}^n} \Xi_i^A X^A. \tag{17}$$

Proof. In view of Theorem 1, the unique solution of the system

$$\bigvee_{A \in \{0,1\}^n} f(A)\Xi_i^A = f(\Xi_i) \qquad (i = 1,\ldots,2^n) \tag{18}$$

for the unknowns $f(A)$ is

$$f(A) = \prod_{i=1}^{2^n} \left[(\Xi_i^A)' \vee f(\Xi_i) \right]$$
$$= \left(\bigvee_{i=1}^{2^n} f'(\Xi_i)\Xi_i^A \right)' = \bigvee_{i=1}^{2^n} f(\Xi_i)\Xi_i^A,$$

where in the last step we have taken into account the orthonormality of the vectors $(\Xi_i^A)_{i=1,\ldots,2^n}$. It follows that

$$\begin{aligned} f(X) &= \bigvee_{A\in\{0,1\}^n} f(A)X^A \\ &= \bigvee_{A\in\{0,1\}^n} \bigvee_{i=1}^{2^n} f(\Xi_i)\Xi_i^A X^A \\ &= \bigvee_{i=1}^{2^n} \bigvee_{A\in\{0,1\}^n} f(\Xi_i)\Xi_i^A X^A \\ &= \bigvee_{i=1}^{2^n} f(\Xi_i) \bigvee_{A\in\{0,1\}^n} \Xi_i^A X^A. \end{aligned}$$

The first author was a participant in the exchange program between the National Academy of Sciences of the U.S.A. and the Academy of the Socialist Republic of Romania.

REFERENCES

1 L. Löwenheim, Gebietdeterminanten, *Math. Ann.* 79:222–236 (1919).

2 W. L. Parker and B. A. Bernstein, On uniquely solvable Boolean equations, *Univ. Calif. Publ. Math., N.S.* 3(1):1–29 (1935).

3 S. Rudeanu, *Boolean Functions and Equations*, American Elsevier, New York, 1974.

4 S. Rudeanu, Systems of linear Boolean equations, *Publ. Inst. Math.* (*Beograd*) (*N.S.*) 22(36):231–235 (1977).

5 S. Rudeanu, Linear Boolean equations and generalized minterms, *Discrete Math.* 43:241–248 (1983).

Received 27 October 1982; revised 2 March 1983