Some Remarks on Boolean Sums*

Kurt Mehlhorn

Fachbereich 10 – Angewandte Mathematik und Informatik, Universität des Saarlandes,
D-6600 Saarbrücken, Germany (Fed. Rep.)

Summary. Neciporuk, Lamagna/Savage and Tarjan determined the monotone network complexity of a set of Boolean sums if any two sums have at most one variable in common. Wegener then solved the case that any two sums have at most $k$ variables in common. We extend his methods and results and consider the case that any set of $h+1$ distinct sums have at most $k$ variables in common. We use our general results to explicitly construct a set of $n$ Boolean sums over $n$ variables whose monotone complexity is of order $n^{5/3}$. The best previously known bound was of order $n^{3/2}$. Related results were obtained independently by Pippenger.

1. Introduction, Notations and Results

We consider the monotone network complexity of sets of Boolean sums $f = (f_1, \ldots, f_m): \{0, 1\}^n \rightarrow \{0, 1\}^m$ with

$$f_i = \bigvee_{j \in F_i} x_j \quad \text{and} \quad F_i \subseteq \{1, \ldots, n\}.$$

Sets of Boolean sums were also considered by Neciporuk, Lamagna/Savage, Tarjan, Wegener and Pippenger.

$C_B(f)$ denotes the network complexity of $f$ over the basis $B$; we will consider $B = \{ \lor \}$ and $B = \{ \lor, \land \}$. A set of Boolean sums is called $(h, k)$-disjoint if for all pairwise distinct $i_0, i_1, i_2, \ldots, i_h$: $|F_{i_0} \cap F_{i_1} \cap \ldots \cap F_{i_h}| \leq k$. It is possible to represent a set of Boolean sums $f: \{0, 1\}^n \rightarrow \{0, 1\}^m$ by a bipartite graph with inputs $\{x_1, \ldots, x_n\}$ and outputs $\{f_1, \ldots, f_m\}$. The edge $(x_j, f_i)$ is present if and only if $j \in F_i$. Then $(h, k)$-disjointness is equivalent to saying that the associated bipartite graph does not contain $K_{k+1,k+1}$ (= complete bipartite graph with $k+1$ inputs and $h+1$ outputs).

* This paper was presented at the MFCS 79 Symposium, Olomouc, Sept. 79
Theorem 1. Let \( f: \{0, 1\}^n \to \{0, 1\}^m \) be a \((h, k)\)-disjoint set of Boolean sums. Then
\[
C_{\lor, \land}(f) \geq \sum_{i=1}^{m} (|F_i|/k-1)/h \cdot \max(1, h-1)
\]

Neciporuk, Lamagna/Savage, Tarjan proved the theorem in the case \( h=1=k \). Wegener extended their results to the case \( h=1 \) and arbitrary \( k \). The first three authors used their result to explicitly construct sets of \( n \) Boolean sums over \( n \) variables whose monotone network complexity is \( \Omega(n^{3/2}) \).

We explicitly construct sets of Boolean sums
\[
f: \{0, 1\}^n \to \{0, 1\}^m
\]
such that \( C_{\lor, \land}(f) = \Omega(n^{5/3}) \). This result was independently obtained by Pippenger.

2. Proofs

Our proof of Theorem 1 is based on two Lemmas. In these Lemmas we will make use of complexity measure \( C^*_B \). \( C^*_B(f) \) is the network complexity of \( f \) over the basis \( B \) under the assumption that all sums \( \bigvee_{x_j} \) with \( |F| \leq k \) are given for free, i.e. the sums \( \bigvee_{x_j} \) can be used as additional inputs.

Measure \( C^*_B \) was introduced by Wegener.

Lemma 1. Let \( f: \{0, 1\}^n \to \{0, 1\}^m \) be a \((h, k)\)-disjoint set of Boolean sums. Then
\[
a) \quad C^*_B(f) \leq \max \{1, h-1\} \cdot C_{\lor, \land}(f), \\
b) \quad C_{\lor, \land}(f) \leq \max \{1, h-1, k-1\} \cdot C_{\lor, \land}(f).
\]

Proof. a) Let \( N \) be an optimal \(*\)-network for \( f \) over the basis \( \{\lor, \land\} \). Then \( N \) contains \( s \lor \)-gates and \( t \land \)-gates, \( s + t = C^*_B(f) \).

For \( i=0, 1, \ldots, t \) we show the existence of a \(*\)-network \( N_i \) for \( f \) with \( \leq t-i \land \)-gates and \( \leq s + (h-1) \cdot i \lor \)-gates.

We have \( N_0 = N \). Suppose now \( N_i \) exists. If \( N_i \) does not contain an \( \land \)-gate then we are done. Otherwise let \( G \) be a last \( \land \)-gate in topological order, i.e. between \( G \) and the outputs there are no other \( \land \)-gates. Let \( g \) be the function computed by \( G \), \( g_1 \) and \( g_2 \) the functions at the input lines of \( G \). Then
\[
g = s_1 \lor \ldots \lor s_p \lor t_1 \lor \ldots \lor t_q,
\]
where \( s_i \) is a variable and \( t_j \) is of length at least 2, is the monotone disjunctive normal form of \( g \).

Case 1: \( p \leq k \). The sum \( s_1 \lor \ldots \lor s_p \) comes for free. By Theorem I of Mehlhorn/Galil \( g \) may be replaced by \( s_1 \lor \ldots \lor s_p \) and an equivalent circuit is obtained.
This shows the existence of network $N_{i+1}$ with $\leq t - i - 1$ $\land$-gates and $\leq s + (h - 1)(i + 1)$ $\lor$-gates.

**Case 2:** $p > k$. There are some outputs, say $f_1, f_2, \ldots, f_l$, depending on $G$. Between $G$ and the output $f_j$ there are only $\lor$-gates and hence $f_j = g \lor u_j$. Since $f_j$ is a boolean sum, $u_j$ is not the constant 1. Hence $\{s_1, \ldots, s_p\} \subseteq F_j$ for $j = 1, \ldots, l$. Since $f$ is $(h, k)$-disjoint we conclude $l \leq h$.

**Claim.** For every $j$, $1 \leq j \leq l$: either $f_j = g_1 \lor u_j$ or $f_j = g_2 \lor u_j$.

**Proof.** Since $g = g_1 \land g_2$ and $f_j = g \lor u_j$ we certainly have $f_j \leq g_1 \lor u_j$ and $f_j \leq g_2 \lor u_j$. Suppose both inequalities are proper. Then there are assignments $\alpha_1, \alpha_2 \in \{0, 1\}^n$ with $f_j(\alpha_1) = 0 < 1 = (g_1 \lor u_j)(\alpha_1)$ and $f_j(\alpha_2) = 0 < 1 = (g_2 \lor u_j)(\alpha_2)$.

Let $\alpha = \max(\alpha_1, \alpha_2)$. Since $f_j$ is a boolean sum $f_j(\alpha) = 0$ and since $g_1 \lor u_j$ and $g_2 \lor u_j$ are monotone $(g_1 \lor u_j)(\alpha) = (g_2 \lor u_j)(\alpha) = 1$. Hence either $u_j(\alpha) = 1$ or $g_1(\alpha) = g_2(\alpha) = 1$ and hence $g(\alpha) = 1$. In either case we conclude $f_j(\alpha) = (g \lor u_j)(\alpha) = 1$.

Contradition. \(\square\)

We obtain circuit $N_{i+1}$ equivalent to $N_i$ as follows.

1) Replace $g$ by the constant 0. This eliminates $\land$-gate $G$ and at least one $\lor$-gate. After this replacement the output line corresponding to $f_j$, $1 \leq j \leq l$, realizes function $u_j$.

2) For every output $f_j$, $1 \leq j \leq l$, we use one $\lor$-gate to sum $u_j$ and $g_1$ (resp. $g_2$). This adds $l \leq h$ $\lor$-gates.

Circuit $N_{i+1}$ has $\leq s + (h - 1)(i + 1)$ $\lor$-gates and $\leq t - i - 1$ $\land$-gates.

In either case we showed the existence of $*$-network $N_{i+1}$. Hence there exists a $*$-network realizing $f$ and containing at most $s + (h - 1) \cdot t \leq \max\{1, h - 1\} \cdot (s + t)$ $\lor$-gates and no $\land$-gates. This ends the proof of part a.

b) In order to prove b) we only have to observe that in case 1) above (i.e. $p \leq k$) we can explicitly compute $s_1 \lor \ldots \lor s_p$ using at most $k - 1$ $\lor$-gates. Hence $N_{i+1}$ contains at most $(k - 1)$ additional $\lor$-gates. \(\square\)

Lemma 1 has several interesting consequences. Firstly it shows that $\land$-gates can reduce the monotone network complexity of sets of $(h, k)$-disjoint Boolean sums by at most a constant factor. Secondly, the proof of Lemma 1 shows that optimal circuits for $(1, 1)$-disjoint sums use no $\land$-gates and that there is always an optimal monotone circuit for $(2, 2)$-disjoint sums without any $\land$-gates.

**Lemma 2.** Let $f: \{0, 1\}^n \rightarrow \{0, 1\}^m$ be a $(h, k)$-disjoint set of Boolean sums. Then

$$C_\lor(f) \geq C_\lor^*(f) \geq \sum_{i=1}^m \binom{|F_i|}{|F_i|/k^\land - 1}/h.$$ 

**Proof.** Let $S$ be an optimal $*$-network over the basis $B = \{\lor\}$. Since $f_i = \lor_{j \in F_i} x_j$ and input lines represent sums of at most $k$ variables output $f_i$ is connected to at least $|F_i|/k^\land$ inputs.

Let $G$ be any gate in $S$. Since $S$ is optimal $G$ realizes a sum of $>k$ variables
and hence at most \( h \) outputs \( f_i \) depend on \( G \) (cf. the discussion of case 2 in the proof of Lemma 1).

For every gate \( G \) let \( n(G) \) be the number of outputs \( f_i \) depending on \( G \). Then \( n(G) \leq h \) and hence

\[
\sum_{G \in S} n(G) \leq h \cdot C_{v,\lambda}^*(f).
\]

Next consider the set of all gates \( H \) connected to output \( f_i \), \( 1 \leq i \leq m \). This subcircuit must contain a binary tree with \( \lceil |F_i|/k^3 \rceil \) leaves, (corresponding to the input lines connected to \( f_i \)) and hence contains at least \( \lceil |F_i|/k^3 \rceil - 1 \) gates. This shows

\[
\sum_{G \in S} n(G) = \sum_{i=1}^{m} \text{number of gates connected to output } f_i
\geq \sum_{i=1}^{m} \left( \lceil |F_i|/k^3 \rceil - 1 \right).
\]

\[\square\]

Wegener proved Lemmas 1 and 2 for the case \( h = 1 \). This special case is considerably simpler to prove. Pippenger proved Lemma 2 by a more complicated graph-theoretic approach.

Theorem 1 is now an immediate consequence of Lemmas 1 and 2. Namely,

\[
C_{v,\lambda}(f) \geq C_{v,\lambda}^*(f) \geq C_{v,\lambda}^*(f)/\max(1, h-1) \geq \sum_{i=1}^{m} (|F_i|/k-1)/h \cdot \max(1, h-1)
\]

by definition of \( C_{v,\lambda}^* \), by Lemma 1a, and by Lemma 2.

3. Explicit Construction of a "Hard" Set of Boolean Sums

Brown exhibited bipartite graphs with \( n \) inputs and outputs, \( \Omega(n^{5/3}) \) edges, and containing no \( K_{3,3} \).

His construction is as follows. Let \( p \) be an odd prime and let \( d \) be a non-zero element of \( GF(p) \) (the field of integers modulo \( p \)), such that \( d \) is a quadratic non-residue modulo \( p \) if \( p \equiv 1 \mod 4 \), and a quadratic residue modulo \( p \) if \( p \equiv 3 \mod 4 \). Let \( H \) be a bipartite graph with \( n = p^3 \) inputs and outputs. The inputs (and outputs) are the triples \( (a_1, a_2, a_3) \) with \( a_1, a_2, a_3 \in GF(p) \). Input \( (a_1, a_2, a_3) \) is connected to output \( (b_1, b_2, b_3) \) if

\[(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^3 = d \mod p.
\]

Brown has shown that bipartite graph \( H \) has \( p^4(p-1) \) edges and that it contains no copy of \( K_{3,3} \).

By the remark in the introduction a bipartite graph corresponds in a natural way to a set of boolean sums. Here we obtain a set of boolean sums over \( \{x_1, \ldots, x_n\} \) with \( \sum_{i=1}^{n} |F_i| = \Omega(n^{5/3}) \).
Furthermore, this set of boolean sums is (2,2)-disjoint. Theorem 1 implies that the monotone complexity of this set of boolean sums is $\Omega(n^{5/3})$.

References

Wegener, I.: A new lower bound on the monotone network complexity of boolean sums, Preprint, Dept. of Mathematics, University of Bielefeld, 1978

Received November 1978 / Revised April 24, 1979