Electronic Colloquium on Computational Complexity, Report No. 49 (2007)

Exact OBDD Bounds for some Fundamental Functions

Beate Bollig, Niko Range, and Ingo Wegener

FB Informatik, LS2, Univ. Dortmund, 44221 Dortmund, Germany
beate.bollig@uni-dortmund.de, niko@range.at, ingo.wegener@uni-dortmund.de

Abstract. Ordered binary decision diagrams (OBDDs) are nowadays the most common dynamic data structure or representation type for Boolean functions. Among the many areas of application are verification, model checking, computer aided design, relational algebra, and symbolic graph algorithms. Although many even exponential lower bounds on the OBDD size of Boolean functions are known, there are only few functions where the OBDD size is even asymptotically known exactly. In this paper the exact OBDD sizes of the fundamental functions multiplexer and addition of n-bit numbers are determined.

1 Introduction and Results

When working with Boolean functions as in circuit verification, synthesis, model checking, and even in graph algorithms, ordered binary decision diagrams, denoted OBDDs, introduced by Bryant (1986), are the most often used data structure supporting all fundamental operations on boolean functions.

Definition 1. *Let $X_n = \{x_1, \dots, x_n\}$ be a set of Boolean variables. A variable order π on X_n is a permutation on $\{1, \dots, n\}$ leading to the ordered list $x_{\pi(1)}, \dots, x_{\pi(n)}$ of the variables.*

Definition 2. *A π-OBDD on X_n (see Figure 1) is a directed acyclic graph $G = (V, E)$ whose sinks are labeled by Boolean constants and whose non sink (or inner) nodes are labeled by Boolean variables from X_n. Each inner node has two outgoing edges one labeled by 0 and the other by 1. The edges between inner nodes have to respect the variable order π, i.e., if an edge leads from an x_i-node to an x_j-node, $\pi^{-1}(i) \leq \pi^{-1}(j)$ (x_i precedes x_j in $x_{\pi(1)}, \dots, x_{\pi(n)}$). Each node v represents a Boolean function $f_v : \{0,1\}^n \to \{0,1\}$ defined in the following way. In order to evaluate $f_v(a)$, $a \in \{0,1\}^n$, start at v. After reaching an x_i-node choose the outgoing edge with label a_i until a sink is reached. The label of this sink defines $f_v(a)$. The size of the π-OBDD G is equal to the number of its nodes.*

Note, that OBDDs are not restricted to the representation of single-output functions. An OBDD represents a Boolean function $f \in B_{n,m}$ by representing simultaneously the outputs $f_1, f_2, \dots, f_m$ of f.

ISSN 1433-8092

The size of the reduced π-OBDD representing f is described by the following structure theorem (Sieling and Wegener (1993)). In order to simplify the description we describe the theorem only for the special case where π equals the identity $id(i) = i$.

Theorem 1. *The number of x_i-nodes of the id-OBDD for $f = (f_1, \ldots, f_m)$ is the number s_i of different subfunctions $f_{j|x_1=a_1,\ldots,x_{i-1}=a_{i-1}}$, $1 \leq j \leq m$ and $a_1, \ldots, a_{i-1} \in \{0,1\}$, essentially depending on x_i (a function g depends essentially on x_i if $g_{|x_i=0} \neq g_{|x_i=1}$).*

The variable order π is not given in advance and we have the freedom (and the problem) to choose a good or even an optimal order for the representation of f. Let π-OBDD(f) denote the π-OBDD size of f.

Definition 3. *The OBDD size of f (denoted by OBDD(f)) is the minimum of all π-OBDD(f).*

It is an obvious aim to determine OBDD(f) for as many of the interesting functions f as exactly as possible. This is similar to other fundamental complexity measures, among them circiut size, formula size, monotone circuit size or algebraic complexity (for such results see Wegener (1987)). Although many even exponential lower bounds on the OBDD size of Boolean functions are known and the method how to obtain such bounds is simple, there are only few functions where the OBDD size is asymptotically known exactly (see, e.g., Bollig and Wegener (2000).) Surprisingly enough, there is only one paper presenting tight bounds on the OBDD size (Wegener (1984)) which has even been published before the notion OBDD was established. For several of the fundamental functions one believes to know the optimal variable order but has no proof for this conjecture. We start to fill this gap by determining exact OBDD bounds for two fundamental functions, namely multiplexer MUX_n, often also called direct storage access function DSA_n, and binary addition ADD_n.

Definition 4. *The multiplexer MUX_n (or direct storage access function DSA_n) is defined on $n + k$ variables $a_{k-1}, \ldots, a_0, x_0, \ldots, x_{n-1}$, where $n = 2^k$. $\mathrm{MUX}_n(a, x) = x_{|a|}$, where $|a|$ is the number whose binary representation equals $(a_{k-1}, \ldots, a_0)$. The a-variables are called address variables and the x-variables data variables.*

Definition 5. *Binary addition ADD_n:$\{0,1\}^{2n} \rightarrow \{0,1\}^{n+1}$ maps two n-bit integers $x = x_{n-1} \ldots x_0$ and $y = y_{n-1} \ldots y_0$ to their sum. That is $\mathrm{ADD}_n(x, y)$= $s_n \ldots s_0$ where $x + y = s$, where $s = s_n \ldots s_0$. $\mathrm{ADD}_{i,n}$ computes the ith bit s_i of ADD_n.*

The results of the paper are the following ones.

Theorem 2. *OBDD*$(\mathrm{MUX}_n) = 2n + 1$.

Theorem 3. *OBDD*$(\mathrm{ADD}_1) = 6$ *and, for $n \geq 2$, OBDD(ADD_n) $= 9n - 5$.*

The upper bounds are contained in Wegener (2000) (Theorem 4.3.2 and Theorem 4.4.3). For binary addition the case $n = 1$ is special since the output of ADD_1 are symmetric functions ($x_0 \wedge y_0$ and $x_0 \oplus y_0$) and the π-OBDD size does not depend on π. Hence, it is sufficient to consider one of the two possible variable orders.

In Sections 2 and 3 the lower bounds are proved where it is essential to avoid an inspection of too many cases since the number of variable orders grows exponentially. The following simple observation will be helpful. Given an arbitrary variable order π the number of nodes labeled by a variable x in the π-OBDD representing a given function f is not smaller than the number of x-nodes in a π-OBDD representing any subfunction of f. Furthermore, the proofs of the lower bounds are based on Theorem 1 implying that we do not introduce a new lower bound method. However, we show how to solve some combinatorial problems in order to obtain more precise results than known before.

2 Tight bounds for the OBDD size of the Multiplexer

In this section, we determine a lower bound on the size of OBDDs for the representation of the multiplexer.

Lemma 1. *The size of an* OBDD *for the representation of the multiplexer is at least* $2n + 1$.

Proof.

Let π be an arbitrary variable order. In order to simplify the description, we assume w.l.o.g. that the sequence of the address variables according to π is $a_0, a_1, \ldots, a_{k-1}$. This assumption is justified because of the observation that the size of an OBDD representing the multiplexer remains the same if we only change the positions of some address variables.

Since the multiplexer depends essentially on all data variables, for each variable x_i, $0 \leq i \leq n-1$, there is at least one node labeled by x_i. Moreover, there have to be two sinks. In the following, our aim is to prove that there exist for each address variable a_i at least 2^i further nodes representing non-constant subfunctions of the multiplexer, such that the number of nodes altogether in the OBDD is at least

$2 + n + \sum_{i=0}^{k-1} 2^i = 2 + n + 2^k - 1 = 2n + 1.$

We fix one of the address variables, called a_i, and use the following notation. Let $T_i(x)$ be the set of the x-variables tested before the variable a_i, $R_i(x)$ describes the set of the remaining x-variables. Now, we consider all possible assignments to the address variables $a_0, \ldots, a_{i-1}$. Our aim is to prove that there exists at least one further node for each assignment. The data variables are partitioned into 2^i disjoint groups such that the indices of the variables of each group agree in their binary representation to the corresponding assignment to the address variables $a_0, a_1, \ldots, a_{i-1}$. Let b_i be an assignment to the address variables $a_0, a_1, \ldots, a_{i-1}$. The group G_{b_i} contains all data variables x_j such that the i least significant bits of the binary representation of j equals b_i. Obviously

$G_{b_{i_1}} \cap G_{b_{i_2}} = \emptyset$ for different assignments b_{i_1} and b_{i_2} to the address variables $a_0, a_1, \ldots, a_{i-1}$.

For each assignment b_i to the address variables we distinguish two cases $a_0, \ldots, a_{i-1}$.

Case 1: $G_{b_i} \cap R_i(x) \neq \emptyset$.

We show that there exists a subfunction corresponding to b_i that essentially depends on a_i, therefore there has to be one further node labeled by a_i in the π-OBDD representing the multiplexer.

For this reason we consider the subfunction which corresponds to the following assignment to the variables. Let x_k be a variable in $G_{b_i} \cap R_i(x)$. The assignment to the address variables $a_0, a_1, \ldots, a_{i-1}$ is b_i, the assignment to all data variables in $T_i(x)$ is 0. Obviously the corresponding subfunction depends essentially on x_k. Therefore, different assignments to the address variables $a_0, a_1, \ldots, a_{i-1}$ lead to different subfunctions which have to be represented at different nodes in the π-OBDD. Furthermore the considered subfunction depends essentially on a_i, since the assignment 1 to x_k, 0 to all other data variables, and the binary representation of k to the address variables has the function value 1 but changing only the assignment to a_i leads to the function value 0.

Altogether we have shown that there has to be one further a_i-node in the π-OBDD representing the multiplexer.

Case 2: $G_{b_i} \cap R_i(x) = \emptyset$.

This case is more difficult because it is possible that there does not exist a subfunction for which the assignment to the variables $a_0, a_1, \ldots, a_{i-1}$ agrees with b_i and which essentially depends on a_i. We have to inspect this case very carefully in order to guarantee that we count each node of the π-OBDD representing the multiplexer only once.

Let b_i^j, $j \leq i$, be the assignment to the variables $a_0, a_1, \ldots, a_{j-1}$ according to b_i. Let i' be the minimum number in $\{0, \ldots, i\}$ such that $G_{b_i^{i'}} \cap R_{i'}(x) = \emptyset$. Since $G_{b_i} \subseteq G_{b_i^{i'}}$ we know that $G_{b_i} \subseteq T_{i'}(x)$. Now, we consider the assignment $b_i^{i'-1}$ which is unique for b_i. Let x_k be the jth data variable of the set $G_{b_i^{i'}}$ in the sequence according to π. Our aim is to show that there are at least 2^{j-1} x_k-nodes in the π-OBDD representing the multiplexer. For this reason we consider the following 2^{j-1} different assignments to the first variables of the set $G_{b_i^{i'}}$ which are before x_k in the sequence according to π. The address variables are set according to the assignment $b_i^{i'-1}$, the data variables in $T_{i'}(x) \setminus G_{b_i^{i'}}$ are fixed to 0 and for the first $j-1$ variables from $G_{b_i^{i'}}$ according to π we consider all possible assignments. Obviously the corresponding subfunctions essentially depend on x_k. Furthermore, two different assignments to the first $j-1$ data variables of $G_{b_i^{i'}}$ according to π lead to different subfunctions, since each of these data variables can determine the output of the multiplexer. For this we consider the following assignments to the remaining variables. Let $x_{k'}$ be one of the first $j-1$ data variables under consideration. The remaining data variables are set to 0 and the address variables are set to the binary representation of k'. The output of the multiplexer is equal to the assignment of $x_{k'}$.

Using the fact that $|G_{b_i^{i'}}| = 2^{k-(i'-1)}$ we can conclude that there are at least

$$\sum_{j=0}^{2^{k-(i'-1)}} 2^j = 2^{2^{k-i'}+1} - 1$$

nodes labeled by a data variable from $G_{b_i^{i'}}$ in the π-OBDD representing the multiplexer. We have already counted one node for each data variable, therefore we have shown that there are at least

$$(2^{2^{k-i'}+1} - 1) - 2^{k-(i'-1)}$$

further nodes.

On the other hand, there are $2^{\ell-(i'-1)}$ assignments b_ℓ to the address variables $a_0, a_1, \ldots, a_{\ell-1}$, $\ell \geq i'-1$, such that $b_i^{i'-1}$ is equal to $b_\ell^{i'-1}$. Therefore, we can conclude that there are

$$\sum_{\ell=i'-1}^{k-1} 2^{\ell-(i'-1)} = \sum_{\ell=0}^{k-i'} 2^\ell = 2^{k-i'+1} - 1$$

assignments to the address variables corresponding to $b_i^{i'-1}$ that lead to case 2 in our investigation.

Since $2^{2^{k-i'+1}} - 1 \geq 2^{k-i'+2} - 1$, we are done.

□

3 Tight bounds for the OBDD size of Binary Addition

As noted before Wegener (2000) has already presented the upper bound of $9n-5$ on the OBDD size of binary addition for two n-bit numbers where $n \geq 2$. In the following, we prove the matching lower bound. Figure 1 shows an OBDD for the binary addition of two 4-bit numbers according to an optimal variable order.

In order to obtain lower bounds on the size of OBDDs one-way communication complexity has become a standard technique (see Hromkovič (1997) and Kushilevitz and Nisan (1997) for the theory of communication theory). In the following, we do not really use methods from communication theory but the notion of a communication matrix which is nothing else but the value table of a function in a different form. A function $f : \{0,1\}^m \times \{0,1\}^n \to \{0,1\}$ can be described by a matrix of size $2^m \times 2^n$. The matrix entry at position (a, b), $a \in \{0,1\}^m$ and $b \in \{0,1\}^n$, is $f(a, b)$. The number of different rows is equal to the number of different subfunctions obtained by the replacement of the first m variables by constants. Since each column is associated with an assignment to the last n variables, a row corresponds to a subfunction essentially depending on a variable z iff there exist two columns associated with two assignments that differ only in the assignment of z and for which the entries in the matrix are different.

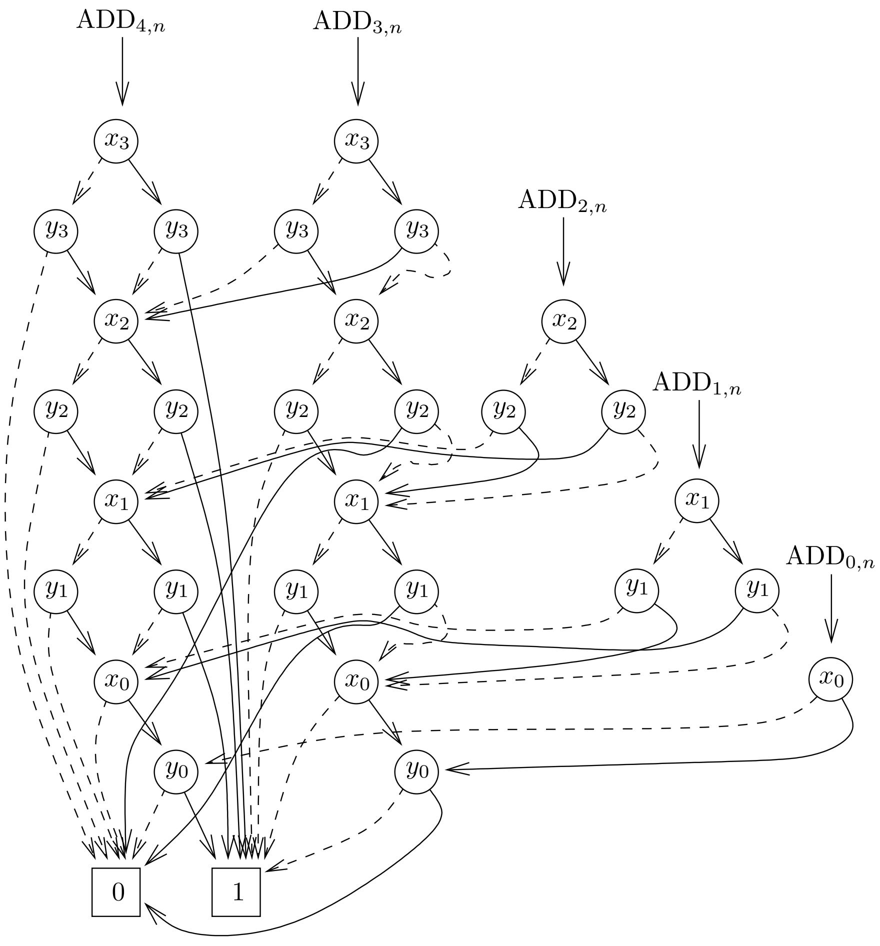

Fig. 1. *An* OBDD *for the binary addition of 4-bit numbers*

Since the functions $\mathrm{ADD}_{i,n}$, $0 \leq i \leq n$, are different and non-constant there are at least $n+1$ nodes representing $\mathrm{ADD}_{i,n}$ in an OBDD representing binary addition. Our aim is to show that for almost all pairs (x_i, y_i), $0 \leq i \leq n-1$, there exist at least 8 nodes labeled by x_i or y_i not representing one of the functions $\mathrm{ADD}_{i,n}$, $0 \leq i \leq n$. Together with the two sinks we are done.

We start our investigation with the following two simple observations. Let π be an arbitrary variable order. Symmetric variables for a given function f are variables that can be exchanged without changing the considered function, i.e. the variables z_i and z_j are symmetric variables for f when $f_{|z_i=0,z_j=1} = f_{|z_i=1,z_j=0}$. In order to simplify the description, we assume w.l.o.g. that for each variable pair (x_i, y_i), $0 \leq i \leq n-1$, the variable x_i ist tested before the variable y_i according to π. This assumption is justified because of the observation that x_i and y_i are symmetric variables for binary addition.

Since the functions $\mathrm{ADD}_{i,n}$, $0 \leq i \leq n-1$, essentially depend on the variables $x_0, y_0, x_1, \ldots, x_i, y_i$ and $\mathrm{ADD}_{n,n}$ essentially depends on all variables, none of the functions $\mathrm{ADD}_{i,n}$, $0 \leq i \leq n$, can be represented at a node labeled by a y-variable.

Now, we introduce some useful notation. Let X be the set of all x-variables and Y the set of all y-variables. The set $X^{>i}$ contains the variables $x_{i+1}, \ldots, x_{n-1}$. Similar the sets $Y^{>i}$, $X^{<i}$, and $Y^{<i}$ are defined.

Let $\Pi_{x_i} = (A_{x_i}, B_{x_i})$, $0 \leq i \leq n-1$, be a partition of the variables in $X \cup Y$ according to a given variable order π, where A_{x_i} contains all variables that are tested before x_i according to π and B_{x_i} the remaining variables. Sometimes we abbreviate (A_{x_i}, B_{x_i}) by (A, B) when the meaning is clear from the context. Similar, Π_{y_i} is defined. For a subset $S \subseteq X \cup Y$, we denote by $\mathcal{A}(S)$ the set of all possible assignments to the variables in S.

Lemma 2. *Let π be an arbitrary variable order and $\Pi_{x_i} = (A_{x_i}, B_{x_i})$ be a partition of the variables in $X \cup Y$ according to π and an arbitrary chosen variable x_i, $0 \leq i \leq n-1$. Let G be a π-OBDD representing* ADD_n. *If $A_{x_i} \cap X^{>i} \neq \emptyset$, the number of x_i-nodes in G not representing one of the functions* $\mathrm{ADD}_{j,n}$, *$0 \leq j \leq n$, is at least* 2.

Proof. Let $h := \min\{j \mid x_j \in A_{x_i} \cap X^{>i}\}$. In the following, we define two assignments a_0 and a_1 to the variables in A_{x_i} and two different assignments b_0 and b_1 to the variables in B_{x_i}. Our aim is to prove that $\mathrm{ADD}_{h,n|a_0}$ and $\mathrm{ADD}_{h,n|a_1}$ are two different subfunctions that essentially depend on x_i. In a_0, a_1, b_0, and b_1, the variables $y_i, \ldots, y_{h-1}$ are set to 1, the remaining y-variables are set to 0. Furthermore, all variables in $X \setminus \{x_h, x_i\}$ are set to 0. In a_0 the variable x_h is fixed to 0, in a_1 to 1. In b_0 the variable x_i is set to 0, in b_1 to 1. Figure 2 illustrates the replacement of some of the variables by constants. The x- and the y-inputs are shown and A and B indicate the set to which the correpondent variable belong.

Table 1 shows a part of the communication matrix for $\mathrm{ADD}_{h,n} : \mathcal{A}(A_{x_i}) \times \mathcal{A}(B_{x_i}) \rightarrow \{0,1\}$. Obviously, the subfunctions $\mathrm{ADD}_{h,n|a_0}$ and $\mathrm{ADD}_{h,n|a_1}$ are different and essentially depend on x_i.

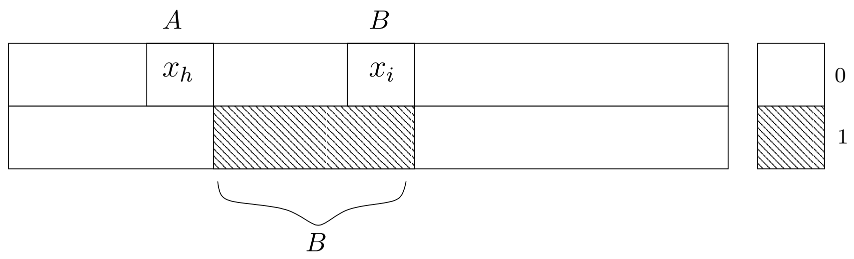

Fig. 2. *The replacement of some of the variables by constants in the proof of Lemma 2*

$\mathrm{ADD}_{h,n}$	b_0	b_1
a_0	0	1
a_1	1	0

Table 1. *Part of the communication matrix for $ADD_{h,n}$*

It remains to prove that the subfunctions $\mathrm{ADD}_{h,n|a_0}$ and $\mathrm{ADD}_{h,n|a_1}$ are different from the functions $\mathrm{ADD}_{0,n}$, ..., $\mathrm{ADD}_{n,n}$. Since $\mathrm{ADD}_{h,n|a_0}$ and $\mathrm{ADD}_{h,n|a_1}$ do not depend on x_h but essentially depend on x_i the subfunctions are different from $\mathrm{ADD}_{j,n}$ for $j \geq h$ or $j < i$. In order to prove that the subfunctions $\mathrm{ADD}_{h,n|a_0}$ and $\mathrm{ADD}_{h,n|a_1}$ are different from the functions $\mathrm{ADD}_{j,n}$ for $i \leq j < h$ we investigate some assignments to the variables in B_{x_i}. Because of the choice of h we know that the variables $x_{i'}$ and $y_{i'}$, $i' \in \{i, \ldots, h-1\}$ are in B_{x_i}. If one of the variables x_{i-1} or y_{i-1} is in A_{x_i}, the subfunctions $\mathrm{ADD}_{h,n|a_0}$ and $\mathrm{ADD}_{h,n|a_1}$ have to be different from the functions $\mathrm{ADD}_{j,n}$, $i \leq j < h$, since the subfunctions do not essentially depend on x_{i-1} and y_{i-1}. Therefore, we can assume that $x_{i-1}, y_{i-1} \in B_{x_i}$. In b'_0 all variables are set to 0, in b'_1 only the variable $y_{i'}$ is set to one. Obviously, the function value for $\mathrm{ADD}_{i',n}$ is 0 for b'_0 and 1 for b'_1. The function value for $\mathrm{ADD}_{h,n|a_0}$ is 0 for b'_0 and b'_1, the function value for $\mathrm{ADD}_{h,n|a_1}$ is 1 for b'_0 and b'_1. Therefore, the two subfunctions $\mathrm{ADD}_{h,n|a_0}$ and $\mathrm{ADD}_{h,n|a_1}$ are different from $\mathrm{ADD}_{i',n}$, $i' \in \{i, \ldots, h-1\}$. □

Lemma 3. *Let π be an arbitrary variable order and $\Pi_{x_i} = (A_{x_i}, B_{x_i})$ be a partition of the variables in $X \cup Y$ according to π and an arbitrary chosen variable x_i, $0 \leq i \leq n-1$. Let G be a π-OBDD representing ADD_n. If $A_{x_i} \cap X^{<i} \neq \emptyset$, the number of x_i-nodes in G not representing one of the functions $\mathrm{ADD}_{j,n}$, $0 \leq j \leq n$, is at least 4.*

Proof. Let x_l be an arbitrary variable in $A_{x_i} \cap X^{<i}$. In the following, we define two assignments a_0 and a_1 to the variables in A_{x_i} that differ only in the assignment to the variable x_l. Our first aim is to prove that $\mathrm{ADD}_{i,n|a_0}$, $\mathrm{ADD}_{i,n|a_1}$, $\mathrm{ADD}_{n,n|a_0}$, and $\mathrm{ADD}_{n,n|a_1}$ are four different subfunctions that essentially depend on x_i. For this reason, we define four assignments b_{00}, b_{01}, b_{10}, and b_{11} to

the variables in B_{x_i} that differ only in the assignments to the variables x_i and y_i.

In a_j, $j \in \{0,1\}$, the variable x_l is set to j. All y-variables in A_{x_i} are set to 1, the remaining x-variables in A_{x_i} to 0. In $b_{j_1 j_2}$, $j_1, j_2 \in \{0,1\}$, the variable x_i is set to j_1, the variable y_i to j_2. The remaining y-variables in B_{x_i} are set to 1, the remaining x-variables to 0. Table 2 shows part of the communication matrix for the functions $\mathrm{ADD}_{i,n}$ and $\mathrm{ADD}_{n,n}$. Obviously, the four subfunctions $\mathrm{ADD}_{i,n|a_0}$, $\mathrm{ADD}_{i,n|a_1}$, $\mathrm{ADD}_{n,n|a_0}$, and $\mathrm{ADD}_{n,n|a_1}$ are different and essentially depending on x_i. Figure 3 illustrates the replacement of the variables with the exception of x_l, x_i, and y_i by constants.

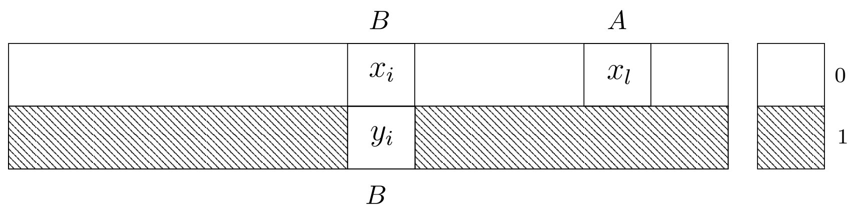

Fig. 3. *The replacement of some of the variables by constants in the proof of Lemma 3*

$\mathrm{ADD}_{i,n}$	b_{00}	b_{01}	b_{10}	b_{11}
a_0	0	1	1	0
a_1	1	0	0	1

$\mathrm{ADD}_{n,n}$	b_{00}	b_{01}	b_{10}	b_{11}
a_0	0	0	0	1
a_1	0	1	1	1

Table 2. *Part of the communication matrix for $ADD_{i,n}$ and $ADD_{n,n}$*

It remains to prove that these four subfunctions are different from the functions $\mathrm{ADD}_{0,n}, \ldots, \mathrm{ADD}_{n,n}$. The functions $\mathrm{ADD}_{0,n}, \ldots, \mathrm{ADD}_{i-1,n}$ do not essentially depend on the variable x_i, whereas the functions $\mathrm{ADD}_{i,n}, \ldots, \mathrm{ADD}_{n,n}$ essentially depend on x_l. Therefore, none of the functions $\mathrm{ADD}_{0,n}, \ldots, \mathrm{ADD}_{n,n}$ is equal to one of the subfunctions $\mathrm{ADD}_{i,n|a_0}$, $\mathrm{ADD}_{i,n|a_1}$, $\mathrm{ADD}_{n,n|a_0}$, and $\mathrm{ADD}_{n,n|a_1}$. $\square$

Combining Lemma 2 and Lemma 3 we obtain the following result.

Corollary 1. *Let π be an arbitrary variable order and let G be a π-OBDD representing ADD_n. The number of x-nodes in G not representing one of the functions $\mathrm{ADD}_{j,n}$, $0 \le j \le n$, is at least $2n-2$.*

Lemma 4. *Let π be an arbitrary variable order and $\Pi_{y_i} = (A_{y_i}, B_{y_i})$ be a partition of the variables in $X \cup Y$ according to π and an arbitrary chosen variable*

y_i, $0 \leq i \leq n-1$. Let G be a π-OBDD representing ADD_n. *If $A_{y_i} \cap X^{>i} \neq \emptyset$ and $B_{y_i} \cap Y^{<i} \neq \emptyset$, the number of y_i-nodes in G is at least* 6.

Proof. Let x_h be a variable in A_{y_i}, $h > i$, and y_l be a variable in B_{y_i}, $l < i$. We consider the following four assignments to the variables in A_{y_i} that differ only in the assignments to the variables x_i and x_h. In $a_{j_1 j_2}$, $j_1, j_2 \in \{0,1\}$, the variable x_i is set to j_1 and the variable x_h to j_2. The remaining x-variables in A_{y_i} are set to 1, the y-variables in A_{y_i} are set to 0. Our aim is to prove that the six subfunctions $\mathrm{ADD}_{h,n|a_{00}}$, $\mathrm{ADD}_{h,n|a_{01}}$, $\mathrm{ADD}_{h,n|a_{10}}$, $\mathrm{ADD}_{h,n|a_{11}}$, $\mathrm{ADD}_{i,n|a_{00}}$, and $\mathrm{ADD}_{i,n|a_{10}}$ are different and that they essentially depend on y_i. For this reason, we consider the following four assignments to the variables in B_{y_i} that differ only in the assignments to the variables y_l and y_i. In $b_{j_1 j_2}$, $j_1, j_2 \in \{0,1\}$, the variable y_l is set to j_1 and the variable y_i to j_2. The x-variables are set to 1 and the remaining y-variables are set to 0. Figure 4 illustrates the replacement of some of the variables by constants.

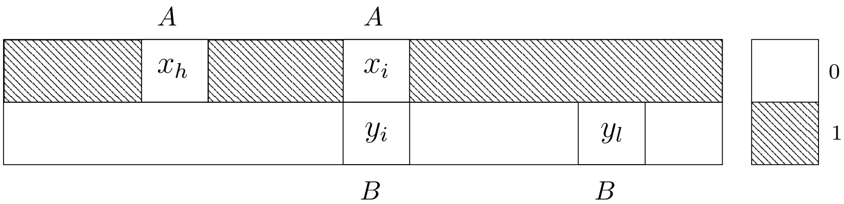

Fig. 4. *The replacement of some of the variables by constants in the proof of Lemma 4*

Table 3 shows part of the communication matrix for the functions $\mathrm{ADD}_{h,n}$ and $\mathrm{ADD}_{i,n}$. Obviously, the six presented subfunctions are different and essentially depend on y_i. Therefore, there are at least 6 y_i-nodes in G.

$\mathrm{ADD}_{h,n}$	b_{00}	b_{01}	b_{10}	b_{11}
a_{00}	0	0	0	1
a_{01}	1	1	1	0
a_{10}	0	1	1	1
a_{11}	1	0	0	0

$\mathrm{ADD}_{i,n}$	b_{00}	b_{01}	b_{10}	b_{11}
a_{00}	0	1	1	0
a_{10}	1	0	0	1

Table 3. *Part of the communication matrix for $ADD_{h,n}$ and $ADD_{i,n}$*

□

Lemma 5. *Let π be an arbitrary variable order and $\Pi_{y_i} = (A_{y_i}, B_{y_i})$ be a partition of the variables in $X \cup Y$ according to π and an arbitrary chosen variable y_i, $0 \leq i \leq n-1$, where $A_{y_i} \cap X^{>i} \neq \emptyset$ and $B_{y_i} \cap Y^{<i} = \emptyset$. Let G be a π-OBDD representing* ADD_n.

i) *The number of y_i-nodes in G is at least 2.*

ii) *If $|B_{y_i} \cap Y| > 1$, there are at least $2 + |B_{y_i} \cap Y|$ nodes labeled by y_i.*

iii) *If $|B_{y_i} \cap Y| \in \{2, 3\}$ and $|B_{y_i} \cap X| < |B_{y_i} \cap Y| - 1$, the number of y_i-nodes in G is at least 6.*

iv) *Let $|B_{y_i} \cap Y| = 3$ and let y_r be the variable in B_{y_i}, where $B_{y_r} \subset B_{y_i}$ and $|B_{y_r}| > 1$, with other words the variable y_i ist tested before y_r but y_r is not the last variable according to π. If $|B_{y_i} \cap X| = |B_{y_i} \cap Y| - 1 = 2$, there are at least 4 y_r-nodes in G and 4 x_r-nodes not representing one of the functions $\mathrm{ADD}_{j,n}$, $0 \le j \le n$.*

Proof. *i)*
Similar to the proofs of Lemma 2 - 4, we show that there are two different subfunctions $\mathrm{ADD}_{i,n|a_0}$ and $\mathrm{ADD}_{i,n|a_1}$ that essentially depend on y_i. We consider the following two assignments to the variables in A_{y_i} that differ only in the assignment to the variable x_i. In a_j, $j \in \{0, 1\}$, the variable x_i is set to j, the remaining x-variables are set to 1, the y-variables are set to 0. Next, we consider two assignments to the variables in B_{y_i} that only differ in the assignment to y_i. In b_j, $j \in \{0, 1\}$, the variable y_i is set to j, the remaining x-variables are set to 1, the y-variables are set to 0. Table 4 shows the corresponding part of the communication matrix for the function $\mathrm{ADD}_{i,n}$. Obviously, the two presented subfunctions are different and essentially depend on y_i. Therefore, there are at least 2 y_i-nodes in G.

$\mathrm{ADD}_{i,n}$	b_0	b_1
a_0	0	1
a_1	1	0

Table 4. *Part of the communication matrix for $ADD_{i,n}$*

ii)
Since $|B_{y_i} \cap Y| > 1$ and $B_{y_i} \cap Y^{<i} = \emptyset$, we can conclude that there exists a variable in $B_{y_i} \cap Y^{>i}$. Let y_h be an arbitrary variable in $B_{y_i} \cap Y^{>i}$. Now, we investigate the subfunctions $\mathrm{ADD}_{h,n|a_1}$ and $\mathrm{ADD}_{n,n|a_1}$. Table 5 shows part of the communication matrix for the functions $\mathrm{ADD}_{h,n}$ and $\mathrm{ADD}_{n,n}$. Obviously, the presented subfunctions are different and essentially depend on y_i.

$\mathrm{ADD}_{h,n}$	b_0	b_1
a_1	1	0

$\mathrm{ADD}_{n,n}$	b_0	b_1
a_1	0	1

Table 5. *Part of the communication matrix for $ADD_{h,n}$ and $ADD_{n,n}$*

It is not difficult to prove that the subfunctions $\mathrm{ADD}_{h,n}$ and $\mathrm{ADD}_{n,n}$ essentially depend on y_h. Since the subfunctions of $\mathrm{ADD}_{i,n}$ do not essentially depend on variables in $Y^{>i}$, we can conclude that the considered subfunctions are different from subfunctions of $\mathrm{ADD}_{i,n}$. Summarizing, we have shown that there are at least $|B_{y_i} \cap Y|$ further y_i-nodes in G.

iii)
Since $B_{y_i} \cap Y^{<i} = \emptyset$ and $|B_{y_i} \cap X| < |B_{y_i} \cap Y| - 1$, there exist a variable $x_h \in A_{y_i}$ and a corresponding variable $y_h \in B_{y_i}$, where $h > i$.

We consider three assignments to the variables in A_{y_i} that differ only in the assignments to the variables x_h and x_i. In a_{01} the variable x_h is set to 0, the variable x_i to 1, in a_{10} the variable x_h is set to 1, the variable x_i to 0, and in a_{11} both are set to 1. The remaining x-variables in A_{y_i} are set to 1, the y-variables in A_{y_i} are set to 0. Our aim is to prove that the subfunctions $\mathrm{ADD}_{i,n|a_{10}}$, $\mathrm{ADD}_{i,n|a_{11}}$, $\mathrm{ADD}_{h,n|a_{01}}$, $\mathrm{ADD}_{h,n|a_{11}}$, $\mathrm{ADD}_{n,n|a_{01}}$, and $\mathrm{ADD}_{i,n|a_{11}}$ are different and that they essentially depend on y_i. For this reason, we consider the following four assignments to the variables in B_{y_i} that differ only in the assignments to the variables y_h and y_i. In $b_{j_1 j_2}$, $j_1, j_2 \in \{0, 1\}$, the variable y_h is set to j_1, the variable y_i to j_2. The x-variables are set to 1 and the remaining y-variables are set to 0. Figure 5 illustrates the replacement of some of the variables by constants.

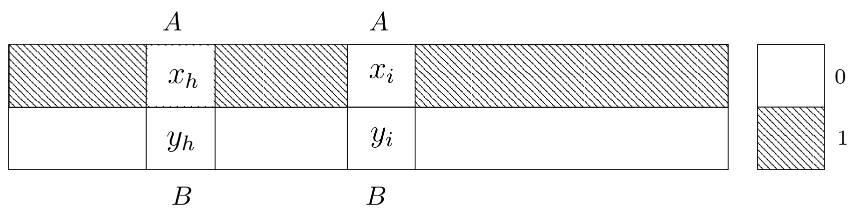

Fig. 5. *The replacement of some of the variables by constants in the proof of Lemma 5 iii)*

Table 6 shows part of the communication matrix for the functions $\mathrm{ADD}_{i,n}$, $\mathrm{ADD}_{h,n}$, and $\mathrm{ADD}_{n,n}$. Obviously, the six presented subfunctions essentially depend on y_i and the subfunctions of $\mathrm{ADD}_{h,n}$ are different from the subfunctions of $\mathrm{ADD}_{n,n}$. Since the subfunctions of $\mathrm{ADD}_{i,n}$ do not essentially depend on y_h, they are different from the other subfunctions. Therefore, there are at least 6 y_i-nodes in G.

iv)
Since $|B_{y_i} \cap X| = |B_{y_i} \cap Y| - 1 = 2$, we know that $x_r \in B_{y_i}$. Using the fact that $x_i \in A_{y_i}$, we can apply Lemma 3 in order to obtain the result that there are at least 4 nodes labeled by x_r in G not representing one of the functions $\mathrm{ADD}_{j,n}$, $0 \le j \le n$.

Let y_z be the variable for which $|B_{y_z}| = 1$, with other words y_z is the last variable according to π.

$\mathrm{ADD}_{i,n}$	b_{00}	b_{01}
a_{10}	0	1
a_{11}	1	0

$\mathrm{ADD}_{h,n}$	b_{00}	b_{01}	b_{10}	b_{11}
a_{01}	0	1	1	0
a_{11}	1	0	0	1

$\mathrm{ADD}_{n,n}$	b_{00}	b_{01}	b_{10}	b_{11}
a_{01}	0	0	0	1
a_{11}	0	1	1	1

Table 6. *Part of the communication matrix for $ADD_{i,n}$, $ADD_{h,n}$, and $ADD_{n,n}$*

Case 1: $r > z$
In a_j the variable x_r is set to j, the remaining x-variables in A_{y_r} are set to 1, the y-variables are set to 0. In b_{j_1,j_2}, $j_1, j_2 \in \{0,1\}$, the variable y_z is set to j_1, the variable y_r to j_2. (If $x_z \in B_{y_r}$, the variable is set to 1.)

Figure 6 illustrates the replacement of some of the variables by constants.

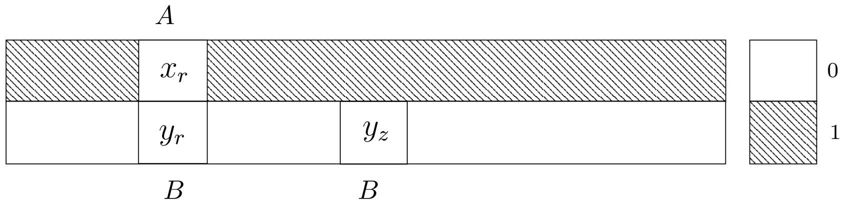

Fig. 6. *The replacement of some of the variables by constants in the proof of Lemma 5 iv)*

Table 7 shows part of the communication matrix for $\mathrm{ADD}_{r,n}$ and $\mathrm{ADD}_{n,n}$. Obviously, the four subfunctions $\mathrm{ADD}_{r,n|a_0}$, $\mathrm{ADD}_{r,n|a_1}$, $\mathrm{ADD}_{n,n|a_0}$, and $\mathrm{ADD}_{n,n|a_1}$ are different and essentially depend on y_r.

$\mathrm{ADD}_{r,n}$	b_{00}	b_{01}	b_{10}	b_{11}
a_0	0	1	1	0
a_1	1	0	0	1

$\mathrm{ADD}_{n,n}$	b_{00}	b_{01}	b_{10}	b_{11}
a_0	0	0	0	1
a_1	0	1	1	1

Table 7. *Part of the communication matrix for $ADD_{r,n}$ and $ADD_{n,n}$*

Case 2: $r < z$
In a_j the variable x_r is set to j, the remaining x-variables in A_{y_r} are set to 1, the y-variables are set to 0. In b_j the variable y_r is set to j. The variable y_z is set to 0. (If $x_z \in B_{y_r}$, the variable is set to 1.)

Figure 7 illustrates the replacement of some of the variables by constants.

Table 8 shows part of the communication matrix for $\mathrm{ADD}_{r,n}$, $\mathrm{ADD}_{z,n}$, and $\mathrm{ADD}_{n,n}$. Obviously, the subfunctions of $\mathrm{ADD}_{z,n}$ and $\mathrm{ADD}_{n,n}$ are different and

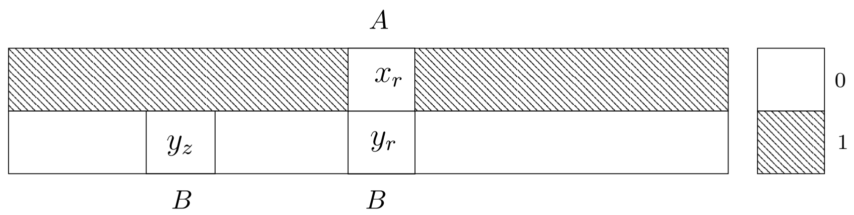

Fig. 7. *The replacement of some of the variables by constants in the proof of Lemma 5 iv)*

essentially depend on y_r. The subfunction of $\mathrm{ADD}_{r,n}$ does not essentially depend on y_z and therefore is different from the other subfunctions.

$\mathrm{ADD}_{r,n}$	b_0	b_1
a_0	0	1
a_1	1	0

$\mathrm{ADD}_{z,n}$	b_0	b_1
a_1	1	0

$\mathrm{ADD}_{n,n}$	b_0	b_1
a_1	0	1

Table 8. *Part of the communication matrix for $ADD_{r,n}$, $ADD_{z,n}$, and $ADD_{n,n}$*

□

In the following, we show that for almost all pairs (x_i, y_i), $0 \le i \le n-1$, the number of nodes not representing one of the functions $\mathrm{ADD}_{j,n}$, $0 \le j \le n$, and labeled by x_i or y_i is 8 if $A_{y_i} \cap X^{>i} = \emptyset$.

Lemma 6. *Let π be an arbitrary variable order and $\Pi_{y_i} = (A_{y_i}, B_{y_i})$ be a partition of the variables in $X \cup Y$ according to π and an arbitrary chosen variable y_i, $0 \le i \le n-1$, where $A_{y_i} \cap X^{>i} = \emptyset$. Let G be a π-OBDD representing* ADD_n.

i) If $A_{x_i} \neq \emptyset$, the number of x_i-nodes in G not representing one of the functions $ADD_{j,n}$, $0 \le j \le n$, is at least 4.
ii) The number of y_i-nodes in G is at least 2.
iii) If $i \le n-2$, the number of y_i-nodes in G is at least 4.
iv) If $|B_{y_i} \cap Y| > 1$, the number of y_i-nodes in G is at least 4.

Table 9 illustrates the minimal number of nodes not representing one of the functions $\mathrm{ADD}_{j,n}$, $0 \le j \le n$, and labeled by x_i or y_i if $A_{y_i} \cap X^{>i} = \emptyset$.

Proof. *i)* Since $A_{x_i} \neq \emptyset$ and $A_{y_i} \cap X^{>i} = \emptyset$, we know that there exist a variable x_k before x_i according to π where $k < i$. Therefore, we can apply Lemma 3 and obtain at least 4 x_i-nodes not representing one of the functions $\mathrm{ADD}_{j,n}$, $0 \le j \le n$.

ii) We consider two assignments to the variables in A_{y_i} that differ only in the assignment to the variable x_i. In a_j, $j \in \{0, 1\}$, the variable x_i is set to j, the

	$A_{x_i} = \emptyset$	$A_{x_i} \neq \emptyset$
$i = n-1 \wedge \lvert B_{y_i} \cap Y \rvert = 1$	2	6
$i \leq n-2 \vee \lvert B_{y_i} \cap Y \rvert > 1$	4	8

Table 9. *The minimal number of x_i- and y_i-nodes if $A_{y_i} \cap X^{>i} = \emptyset$*

remaining x-variables in A_{y_i} are set to 1, the y-variables in A_{y_i} are set to 0. Our aim is to prove that the subfunctions $\mathrm{ADD}_{i,n|a_0}$ and $\mathrm{ADD}_{i,n|a_1}$ are different and that they essentially depend on y_i. For this reason, we consider the following two assignments to the variables in B_{y_i} that differ only in the assignment to the variable y_i. In b_j, $j \in \{0,1\}$, the variable y_i is set to j, the x-variables in B_{y_i} are set to 1 and the remaining y-variables are set to 0. Table 10 show part of the communication matrix for $\mathrm{ADD}_{i,n}$. Obviously, the two subfunctions are different and essentially depend on y_i.

$\mathrm{ADD}_{i,n}$	b_0	b_1
a_0	0	1
a_1	1	0

Table 10. *Part of the communication matrix for $ADD_{i,n}$*

iii) Using part *ii)* it remains to prove that there are two further nodes labeled by y_i. Since $A_{y_i} \cap X^{>i} = \emptyset$, we can conclude that there exist a variable y_h, $h > i$, in $B_{y_i} \cap Y$. Now, we consider the subfunctions $\mathrm{ADD}_{h,n|a_1}$ and $\mathrm{ADD}_{n,n|a_1}$ for a_1 chosen as in part *ii)*. Both essentially depend on y_h and are therefore different from the subfunctions considered in part *ii)*. Table 11 shows part of the communication matrix for $\mathrm{ADD}_{h,n}$ and $\mathrm{ADD}_{n,n}$. Obviously, both subfunctions

$\mathrm{ADD}_{h,n}$	b_0	b_1
a_1	1	0

$\mathrm{ADD}_{n,n}$	b_0	b_1
a_1	0	1

Table 11. *Part of the communication matrix for $ADD_{h,n}$ and $ADD_{n,n}$*

are different and essentially depend on y_i. Together with the proof of part *ii)* we obtain at least 4 y_i-nodes.

iv) We assume $i = n-1$, otherwise we can use part *iii)* and we are done. As in part *ii)* we consider the assignments a_0 and a_1. Using the proof of part *ii)* we know that there have to be 2 y_{n-1}-nodes representing the subfunctions

$\mathrm{ADD}_{n-1,n|a_0}$ and $\mathrm{ADD}_{n-1,n|a_1}$. Our aim is to prove that there have to be two further y_{n-1}-nodes in G representing the subfunctions $\mathrm{ADD}_{n,n|a_0}$ and $\mathrm{ADD}_{n,n|a_1}$. Since $|B_{y_{n-1}} \cap Y| > 1$, there has to be a variable y_l in $B_{y_{n-1}}$, where $l < n-1$. We consider the following four assignments to the variables in $B_{y_{n-1}}$ that differ only in the assignments to the variables y_l and y_{n-1}. In $b_{j_1 j_2}$, $j_1, j_2 \in \{0,1\}$, the variable y_l is set to j_1 and the variable y_{n-1} to j_2. The x-variables are set to 1 and the remaining y-variables are set to 0.

Figure 8 illustrates the replacement of the variables with the exception of x_{n-1}, y_{n-1}, and y_l by constants.

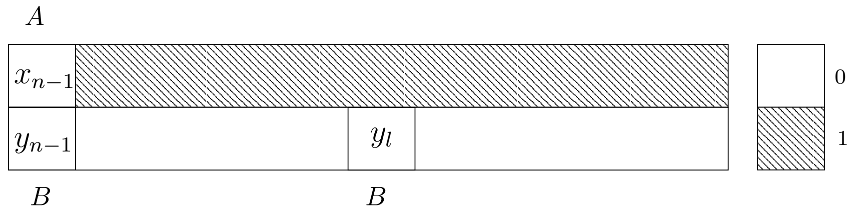

Fig. 8. *The replacement of some of the variables by constants in the proof of Lemma 6*

Table 12 shows part of the communication matrix for $\mathrm{ADD}_{n-1,n}$ and $\mathrm{ADD}_{n,n}$.

$\mathrm{ADD}_{n-1,n}$	b_{00}	b_{01}	b_{10}	b_{11}
a_0	0	1	1	0
a_1	1	0	0	1

$\mathrm{ADD}_{n,n}$	b_{00}	b_{01}	b_{10}	b_{11}
a_0	0	0	0	1
a_1	0	1	1	1

Table 12. *Part of the communication matrix for $ADD_{n-1,n}$ and $ADD_{n,n}$*

Obviously, the four subfunctions are different and essentially depend on y_{n-1}, therefore there are at least 4 nodes labeled by y_{n-1} in G. $\square$

Theorem 4. *The size of an* OBDD *for the representation of binary addition is at least $9n-5$ for $n \geq 2$.*

Proof. Let π be an arbitrary variable order and G be a π-OBDD representing ADD_n. Our aim is to prove that G has at least $9n-5$ nodes. There are $(n+1)+2$ nodes in G representing the functions $\mathrm{ADD}_{i,n}$, $0 \leq i \leq n$, and the constant functions 0 and 1. Using Corollary 1 we obtain at least $2n-2$ further nodes labeled by an x-variable. If we can prove that there are at least $6n-6$ further nodes, we are done.

Now, we investigate the number of y-nodes in G. For each variable y_i, $0 \leq i \leq n-1$, exactly one of the Lemmas 4-6 can be applied and for almost all variables

y_i it can be proved that G contains at least 6 y_i-nodes. In the following, we book for each variable x_i for which we can prove that there exist at least 4 x_i-nodes in G not representing one of the functions $\mathrm{ADD}_{i,n}$, $0 \leq i \leq n$, 2 nodes by the variable y_i. Next, we look more carefully at the y-variables for which we cannot directly guarantee using Lemma 4-6 that there are at least 6 nodes in G. Figure 9 illustrates all possible cases where the number of nodes accounted for a y-variable can be less than 6.

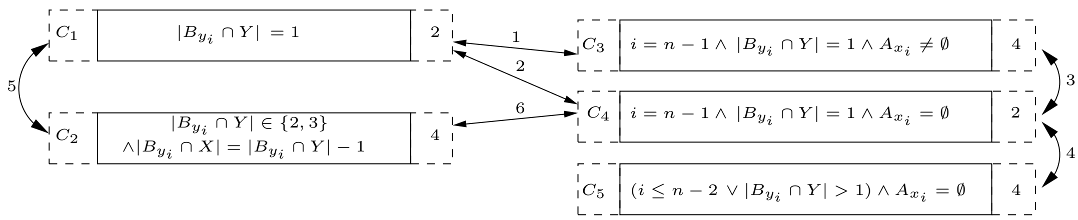

Fig. 9. *Possible sets of y-variables for which less than 6 nodes can be directly booked*

- C_1 is the set of y_i-variables for which $A_{y_i} \cap X^{>i} \neq \emptyset$ and $|B_{y_i} \cap Y| = 1$.
- C_2 is the set of y_i-variables for which $A_{y_i} \cap X^{>i} \neq \emptyset$, $B_{y_i} \cap Y^{<i} = \emptyset$, and $|B_{y_i} \cap Y| \in \{2,3\}$, where $|B_{y_i} \cap X| = |B_{y_i} \cap Y| - 1$.
- C_3 contains the variable y_{n-1} if $|B_{y_{n-1}} \cap Y| = 1$ and $A_{x_{n-1}} \neq \emptyset$.
- C_4 contains the variable y_{n-1} if $|B_{y_{n-1}} \cap Y| = 1$ and $A_{x_{n-1}} = \emptyset$.
- C_5 is the set of y_i-variables, $i \leq n-2$, where $A_{y_i} \cap X^{>i} = \emptyset$, $|B_{y_i} \cap Y| > 1$, and $A_{x_i} = \emptyset$.

The right column of a C_j-row, $1 \leq j \leq 5$, presents the minimal number of nodes accounted for a variable in C_j. Obviously, each variable y_i can be in at most one set C_j, $j \in \{1, \ldots, 5\}$. On the other hand, $|C_j| \leq 1$ for $j \in \{1,3,4,5\}$. Lemma 5 part *iv)* guarantees that C_2 can contain at most one variable. An arrow between a C_{j_1}- and a C_{j_2}-row, $j_1, j_2 \in \{1, 2, \ldots, 5\}$, indicates that there cannot be a variable in C_{j_1} for which only the minimal number of nodes can be accounted for, as well as in C_{j_2}.

It is not difficult to see that $|C_1| + |C_3| + |C_4| \leq 1$, therefore the arrows $1-3$ are justified. Furthermore, using the definition of the sets we can immediatley conclude that $|C_4| + |C_5| \leq 1$ and the arrow 4 follows. For the fifth arrow we have to work a little bit harder.

Claim:
If $C_1 \neq \emptyset$ and $C_2 \neq \emptyset$, the minimal number of nodes accounted for the variables in $C_1 \cup C_2$ is at least 8.

Proof. Let y_i be the variable in C_1 and y_j be the variable in C_2. Because of the definition of C_2 we know that $i > j$ and $x_i \in B_{y_j}$. Therefore, we can apply Lemma 3 in order to prove that there are at least 4 x_i-nodes. □

In order to prove that $|C_2| + |C_4| \leq 1$ and therefore the arrow 6 is justified, we assume that $C_2 \neq \emptyset$. Let $y_j \in C_2$ and $y_k \in B_{y_i}$, with other words the variable y_i is tested before the variable y_k according to π. Because of the definition of C_2, more precisely, since $|B_{y_i} \cap X| = |B_{y_i} \cap Y| - 1$, it follows that also $x_k \in B_{y_i}$ and therefore $A_{x_k} \neq \emptyset$. Therefore, the set C_4 has to be empty.

Summarizing, we obtain the following results:

- $|C_i| \leq 1$, $i \in \{1, 2, \ldots, 5\}$.
- $C_1 \neq \emptyset \Rightarrow C_2 = C_3 = C_4 = \emptyset$.
- $C_4 \neq \emptyset \Rightarrow C_1 = C_2 = C_3 = C_5 = \emptyset$.

Altogether, we have proved that the number of nodes in G accounted for a y-variable is at least $6n - 6$.

$\square$

References

1. Bollig, B. and Wegener, I. (2000). Asymptotically optimal bounds for OBDDs and the solution of some basic OBDD problems. Journal of Computer and System Sciences 61, 558–579.
2. Bryant, R. E. (1986). Graph-based algorithms for Boolean manipulation. IEEE Trans. on Computers 35, 677–691.
3. Hromkovič, J. (1997). *Communication Complexity and Parallel Computing.* Springer.
4. Kushilevitz, E. and Nisan, N. (1997). *Communication Complexity.* Cambridge University Press.
5. Sieling, D. and Wegener, I. (1993). NC-algorithms for operations on binary decision diagrams. Parallel Processing Letters 48, 139–144.
6. Wegener, I. (1984). Optimal decision trees and one-time-only branching-programs for symmetric Boolean functions. Information and Control 62, 129–143.
7. Wegener, I. (1987). *The Complexity of Boolean Functions.* Wiley-Teubner.
8. Wegener, I. (2000). *Branching Programs and Binary Decision Diagrams - Theory and Applications.* SIAM Monographs on Discrete Mathematics and Applications.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/