<table>
<thead>
<tr>
<th>Title</th>
<th>Counterexamples to the long-standing conjecture on the complexity of BDD binary operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Yoshinaka, Ryo; Kawahara, Jun; Denzumi, Shuhei; Arimura, Hiroki; Minato, Shin-ichi</td>
</tr>
<tr>
<td>Citation</td>
<td>Information Processing Letters, 112(16), 636-640</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2012-08-31</td>
</tr>
<tr>
<td>Doc URL</td>
<td><a href="http://hdl.handle.net/2115/50105">http://hdl.handle.net/2115/50105</a></td>
</tr>
<tr>
<td>Type</td>
<td>article (author version)</td>
</tr>
<tr>
<td>File Information</td>
<td>IPL112-16_636-640.pdf</td>
</tr>
<tr>
<td></td>
<td>Hokkaido University Collection of Scholarly and Academic Papers: HUSCAP</td>
</tr>
</tbody>
</table>
Counterexamples to the Long-standing Conjecture on the Complexity of BDD Binary Operations

Ryo Yoshinaka\textsuperscript{a,*}, Jun Kawahara\textsuperscript{a}, Shuhei Denzumi\textsuperscript{b}, Hiroki Arimura\textsuperscript{b}, Shin-ichi Minato\textsuperscript{b,a}

\textsuperscript{a}ERATO MINATO Discrete Structure Manipulation System Project, Japan Science and Technology Agency, Sapporo 060-0814, Japan
\textsuperscript{b}Graduate School of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan

Abstract

In this article, we disprove the long-standing conjecture, proposed by R. E. Bryant in 1986, that his binary decision diagram (BDD) algorithm computes any binary operation on two Boolean functions in linear time in the input-output sizes. We present Boolean functions for which the time required by Bryant’s algorithm is a quadratic of the input-output sizes for all nontrivial binary operations, such as $\land$, $\lor$, and $\oplus$. For the operations $\land$ and $\lor$, we show an even stronger counterexample where the output BDD size is constant, but the computation time is still a quadratic of the input BDD size. In addition, we present experimental results to support our theoretical observations.

Keywords: analysis of algorithms, binary decision diagram, data structures.

1. Introduction

Binary decision diagram (BDD) serves to represent arbitrary Boolean functions as a compact data structure and to efficiently perform various Boolean operations on those functions. Since every Boolean function has a unique BDD representation, we denote the BDD representing Boolean function $f$ by $B(f)$. Bryant proposed an elegant algorithm that performs any Boolean operation on two functions by using BDDs \cite{Bryant86}. Bryant’s algorithm computes $B(f \circ g)$ in $O(|B(f)||B(g)|)$ time for any binary Boolean operation $\circ$ and Boolean functions $f$ and $g$, where $|B(h)|$ denotes the size of the BDD $B(h)$. Moreover, Bryant presented a family of function pairs $f_c$ and $g_c$ for which it takes $\Theta(|B(f_c)||B(g_c)|)$ time to compute $B(f_c \lor g_c)$. Furthermore, Bryant conjectured that his algorithm would run in $O(|B(f)| + |B(g)| + |B(f \circ g)|)$ time. The conjecture has remained open for a quarter of a century. In this article we present a family of Boolean function pairs for which Bryant’s algorithm runs in $\Theta(|B(f)||B(g)|)$ time and $|B(f \circ g)|$ is proportional to $|B(f)| + |B(g)|$. Therefore Bryant’s conjecture does not hold for his algorithm and all the other existing algorithms based on it.

*Corresponding author: ry@i.kyoto-u.ac.jp, Tel.: +81-75-753-5638, Fax.: +81-75-753-5628. Relocated to Kyoto University after this work was completed.
2. Preliminaries

2.1. Binary Decision Diagram

A BDD for representing a Boolean function $f(x_1, \ldots, x_n)$ can be viewed as a labeled directed acyclic graph. We write $x_i < x_j$ if $i < j$. This graph has two special nodes $0$ and $1$ that have no outgoing edges, called terminal nodes. Any other node $p$ is assigned the triple $[x_i, p_0, p_1]$; namely, $p \mapsto [x_i, p_0, p_1]$, where $x_i$ is a variable ($1 \leq i \leq n$) and $p_0$ and $p_1$ are nodes. The assignment $p \mapsto [x_i, p_0, p_1]$ means that $p$ is labeled with the variable $x_i$ and has two outgoing edges to $p_0$ and $p_1$. We name the edge from $p$ to $p_0$ as the $0$-edge of $p$ and that to $p_1$ as the $1$-edge. To make the structure compact and to ensure its uniqueness, the following conditions must be satisfied:

- $p \mapsto [x_i, p_0, p_1], \quad p_0 \mapsto [x_j, q_0, q_1]$, and $p_1 \mapsto [x_k, r_0, r_1]$ implies $x_i < x_j$ and $x_k < x_i$,
- $p \mapsto [x_i, p_0, p_1]$ and $q \mapsto [x_i, p_0, p_1]$ implies $p = q$,
- $p \mapsto [x_i, p_0, p_1]$ implies $p_0 \neq p_1$.

The first and third clauses can be checked locally, whereas this is not the case for the second clause. To ensure that the second clause is satisfied, we maintain a hash table called uniquetable that provides the unique node $p$ such that $p \mapsto [x_i, p_0, p_1]$ (if this exists) for the key $(x_i, p_0, p_1)$.

For the sake of a uniform description of instructions of our algorithms, we conveniently assume that $0 \mapsto [x_{n+1}, 0, 0]$ and $1 \mapsto [x_{n+1}, 1, 1]$, where the associated BDD involves $n$ variables $x_1, \ldots, x_n$. We inductively interpret each node $p$ of a BDD as a Boolean function $\phi_p$ as follows:

- if $p = 0$, $\phi_p$ is the contradiction: $\phi_p = 0$,
- if $p = 1$, $\phi_p$ is the tautology: $\phi_p = 1$,
- if $p$ is not a terminal node and $p \mapsto [x_i, q, r]$, $\phi_p$ is the function of $x_1, \ldots, x_n$ such that\(^{1}\)

\[
\begin{align*}
\phi_p(0, x_{i+1}, \ldots, x_n) &= \phi_q(x_{i+1}, \ldots, x_n) & \text{where } x_j \text{ is the label of } q; \\
\phi_p(1, x_{i+1}, \ldots, x_n) &= \phi_r(x_{i+1}, \ldots, x_n) & \text{where } x_k \text{ is the label of } r.
\end{align*}
\]

Note that the relation between the assigned functions of a node and its children is also known as Shannon’s expansion [4].

We often identify a node $p$ and the BDD consisting of all and only nodes reachable from $p$ as long as no confusion occurs.\(^{2}\) For more details about BDDs, see, for example, [1] or [2].

**Theorem 1** (Theorem 1 in [1]). *For any Boolean function $f$, there exists a unique BDD representing $f$ up to isomorphism.*

---

\(^{1}\)According to conventional interpretation, the sequence $x_i, x_{i+1}, \ldots, x_n$ is regarded as the empty sequence if the initial index $i$ is larger than the final index $n$. In such cases, $\phi_p()$ with an empty argument represents a constant function.

\(^{2}\)We admit $0$ and $1$ as special BDDs with only one terminal node.
We denote the unique BDD for \( f \) by \( B(f) \) and the number of nodes of \( B(f) \) by \( |B(f)| \).

Now, for two functions \( f \) of \( x_1, \ldots, x_n \) and \( g \) of \( x_{k+1}, \ldots, x_n \) with \( k \in \{0, \ldots, n\} \), \( g \) is said to be a \((k\text{-level}) \) subfunction of \( f \) if there are \( c_1, \ldots, c_k \in \{0, 1\} \) such that
\[
g(x_{k+1}, \ldots, x_n) = f(c_1, \ldots, c_k, x_{k+1}, \ldots, x_n).
\]

**Theorem 2** ([2]). The number of nodes of \( B(f) \) is less than or equal to the number of subfunctions of \( f \).

**Lemma 3.** There exist \( 2^{2^n} \) functions of \( m \) variables.

### 2.2. Algorithm for Binary Operations

For a Boolean binary operation \( \circ : \{0, 1\}^2 \rightarrow \{0, 1\} \), we also use the operator \( \circ \) on Boolean function pairs; namely, we define \( f \circ g \) by \( (f \circ g)(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) \circ g(x_1, \ldots, x_n) \). A highly important benefit of using BDDs is their efficient implementation; a binary Boolean operation \( \circ \) can be performed without decompressing the BDDs. Specifically, for any two Boolean functions \( f \) and \( g \) and any binary operation \( \circ \), \( B(f \circ g) \) can be computed from \( B(f) \) and \( B(g) \) in \( O(|B(f)||B(g)|) \) time [1]. Algorithm 1 (Applying) takes a binary operation \( \circ \) and two BDD nodes \( p \) and \( q \), and, by recursively calling itself, computes a node \( r \) such that \( \phi_r = \phi_p \circ \phi_q \). Algorithm 2 (GetNode) is a subroutine called from Algorithm 1.

We assume that \( \circ \) is what we term properly binary; explicitly, neither
- \( a \circ 0 = a \circ 1 \) for any \( a \in \{0, 1\} \) nor
- \( 0 \circ a = 1 \circ a \) for any \( a \in \{0, 1\} \)

holds. Even if \( \circ \) is properly binary, the value of \( \phi_p \circ \phi_q \) is often easily computed in terms of BDD nodes. For example, if \( \circ = \land \), we have that \( \phi_p \land 0 = 0 \) and \( \phi_p \land 1 = \phi_p \).

For inputs \( \langle \land, p, 0 \rangle \) and \( \langle \land, p, 1 \rangle \), the algorithm should immediately return the nodes 0 and \( p \), respectively, without decompressing the function \( \phi_p \). When we can determine \( r \) such that \( \phi_r = \phi_p \circ \phi_q \) without decomposition of the functions, we say that the value of \( \text{Applying}(\circ, p, q) \) is trivial. This covers the case where both \( p \) and \( q \) are terminal nodes. Note that if \( \circ \) is properly binary and neither \( p \) nor \( q \) is a terminal node, then the value of \( \text{Applying}(\circ, p, q) \) is not trivial.

To avoid computing the value for the same pair of arguments twice or more, once the value \( \text{Applying}(\circ, p, q) \) has been computed, we store this value with the key \( \langle p, q \rangle \) to a cache, denoted by \text{cache} in Algorithm 1.

### 3. Construction

#### 3.1. Construction by Multiplexers

Fix a positive integer \( n \) and let \( m = \lceil \log n \rceil \). We assume \( 2n+m \) variables \( x_1, y_1, x_2, y_2, \ldots, x_n, y_n, z_1, \ldots, z_m \) in this order. We then define two functions \( f_n \) and \( g_n \) of these variables by
\[
f_n(x_1, y_1, \ldots, x_n, y_n, z_1, \ldots, z_m) = x_{\beta(z_1, \ldots, z_m)},
\]
\[
g_n(x_1, y_1, \ldots, x_n, y_n, z_1, \ldots, z_m) = y_{\beta(z_1, \ldots, z_m)}.
\]
Algorithm 1 Apply [1]

Input: $\diamond$: binary Boolean operation, $p, q$: BDD nodes
Result: BDD node for $p \diamond q$

1: if the value of $\text{Apply}(\diamond, p, q)$ is trivial then
2:   let $r$ be such that $\phi_r = \phi_p \diamond \phi_q$;
3:   return $r$;
4: else if there is a node $r$ with the key $h_{p \diamond q}$ in cache then
5:   return $r$;
6: else
7:   let $x, p_0, p_1, y, q_0, q_1$ be such that $p \mapsto [x, p_0, p_1]$ and $q \mapsto [y, q_0, q_1]$;
8:   if $x = y$ then
9:     let $r$ be $\text{Getnode}(x, \text{Apply}(\diamond, p_0, q_0), \text{Apply}(\diamond, p_1, q_1))$;
10: else if $x < y$ then
11:     let $r$ be $\text{Getnode}(x, \text{Apply}(\diamond, p_0, q), \text{Apply}(\diamond, p_1, q))$;
12: else
13:     let $r$ be $\text{Getnode}(y, \text{Apply}(\diamond, p, q_0), \text{Apply}(\diamond, p, q_1))$;
14: end if
15: register $r$ to cache with the key $(p, q)$;
16: return $r$;
17: end if

Algorithm 2 Getnode

Input: $x$: variable, $p_0, p_1$: BDD nodes
Result: BDD node with $[x, p_0, p_1]$

1: if $p_0 = p_1$ then
2:   return $p_0$;
3: else if there is a node $p$ with the key $(x, p_0, p_1)$ in uniquetable then
4:   return $p$;
5: else
6:   create a node $p$ with $p \mapsto [x, p_0, p_1]$;
7:   register $p$ to uniquetable with the key $(x, p_0, p_1)$;
8:   return $p$;
9: end if
where
\[
\beta(z_1, \ldots, z_m) = \begin{cases} 
1 + \sum_{k=1}^m 2^{k-1}z_k & \text{if } \sum_{k=1}^m 2^{k-1}z_k < n; \\
1 & \text{otherwise}.
\end{cases}
\]

**Lemma 4.** The numbers of nodes of \(B(f_n)\) and \(B(g_n)\) are at most \(4 \cdot 2^n\), respectively.

**Proof.** We will prove the lemma for \(f_n\), and the same argument then applies to \(g_n\).

\(B(f_n)\) has no nodes labeled \(y_i\) for any \(i\). Therefore, counting the number of other nodes is sufficient, which is bounded by the number of \(\ell\)-level subfunctions of \(f_n\) for \(\ell = 2k\) for \(k = 0, \ldots, n\) and \(\ell = 2n + k\) for \(k = 1, \ldots, m\) by Theorem 2.

An \(\ell\)-level subfunction \(h\) of \(f_n\) for \(\ell = 2k\) with \(k \leq n\) has the form
\[
h(x_0, y_1, \ldots, x_n, y_{n+1}, z_1, \ldots, z_m)
\]
for some \(a_1, \ldots, a_k \in \{0, 1\}\). There are \(2^k\) choices of \(a_1, \ldots, a_k\), and thus the number of \(\ell\)-level subfunctions is at most \(2^k\).

An \(\ell\)-level subfunction \(h\) of \(f_n\) for \(\ell = 2n + k\) with \(1 \leq k \leq m\) is a function of \(m - k\) variables. At most \(2^{2m-1}\) such functions exist by Lemma 3.

Hence,
\[
|B(f_n)| \leq \sum_{k=0}^n 2^k + \sum_{k=1}^m 2^{2m-k} \leq 2 \cdot 2^n + 2 \cdot 2^{2m-1} \leq 4 \cdot 2^n,
\]
where \(2^{m-1} = 2^\lceil \log n \rceil - 1 \leq n\).

\[\square\]

**Lemma 5.** The function \(f_n\) has \(2^n\) distinct \(2n\)-level subfunctions.

**Proof.** For each \(\bar{a} = (a_1, \ldots, a_n) \in \{0, 1\}^n\),
\[
f_{\bar{a},\bar{b}}(z_1, \ldots, z_m) = a_{\beta(z_1, \ldots, z_m)}
\]
is a \(2n\)-level subfunction of \(f_n\). For each \(\bar{a}, \bar{b} \in \{0, 1\}^n\), if \(\bar{a} \neq \bar{b}\), then \(f_{\bar{a},\bar{b}}(z_1, \ldots, z_m)\) and \(f_{\bar{a}',\bar{b}'}(z_1, \ldots, z_m)\) are distinct subfunctions because there exist \(d \in \{1, \ldots, n\}\) and \(c_1, \ldots, c_m \in \{0, 1\}\) such that \(d = \beta(c_1, \ldots, c_m)\) and \(a_d \neq b_d\).

\[\square\]

**Lemma 6.** For any properly binary operation \(\circ\), Algorithm 1 calls itself recursively at least \((2^n)^2\) times with input \((\circ, B(f_n), B(g_n))\).

**Proof.** For each \(\bar{a} \in \{0, 1\}^n\), let \(p_{\bar{a}}\) denote the unique node of \(B(f_n)\) such that
\[
\phi_{p_{\bar{a}}}(z_{k+1}, \ldots, z_m) = f_n(a_1, 0, a_2, 0, \ldots, a_n, 0, c_{k+1}, \ldots, z_m)
\]
for any \(c_1, \ldots, c_k \in \{0, 1\}\), where \(k\) is some natural number. Similarly, let \(q_{\bar{b}}\) denote the unique node of \(B(g_n)\) such that
\[
\phi_{q_{\bar{b}}}(z_{k+1}, \ldots, z_m) = g_n(0, b_1, 0, b_2, \ldots, 0, b_n, c_{k+1}, \ldots, z_m)
\]
for any \(c_1, \ldots, c_k \in \{0, 1\}\), where \(k\) is some natural number. For each pair of \(\bar{a}\) and \(\bar{b} \in \{0, 1\}^n\), the algorithm calls Apply(\(\circ, p_{\bar{a}}, q_{\bar{b}}\)). By Lemma 5, \(p_{\bar{a}}\) and \(q_{\bar{b}}\) are distinct if \(\bar{a} \neq \bar{a}'\) for \(\bar{a}, \bar{a}' \in \{0, 1\}^n\). The same argument holds true for \(q_{\bar{b}}\) and \(q_{\bar{b}'}\), with \(\bar{b} \neq \bar{b}'\).

Hence, the algorithm calls itself at least \((2^n)^2\) times.

\[\square\]
Lemma 7. \( f_n \circ g_n \) has at most \( 6 \cdot 2^n \) subfunctions for any binary operation \( \circ \).

Proof. For each \( \ell \in \{0, \ldots, 2n + m\} \), we count the number of \( \ell \)-level subfunctions. Let \( h \) be a subfunction of

\[
(f_n \circ g_n)(x_1, \ldots, z_m) = x_{\beta(z_1, \ldots, z_m)} \circ y_{\beta(z_1, \ldots, z_m)}.
\]

We have three cases:

Case 1: \( \ell = 2k \) for \( k \in \{0, \ldots, n\} \). There exist \( a_1, b_1, \ldots, a_k, b_k \in \{0, 1\} \) such that

\[
h(x_{k+1}, y_{k+1}, \ldots, z_m) = (f_n \circ g_n)(a_1, b_1, \ldots, a_k, b_k, x_{k+1}, y_{k+1}, \ldots, z_m)
\]

\[
= \begin{cases} 
  c_i & \text{if } \beta(z_1, \ldots, z_m) = i \leq k; \\
  x_i \circ y_i & \text{if } \beta(z_1, \ldots, z_m) = i > k,
\end{cases}
\]

where \( c_i = a_i \circ b_i \) for \( i \in \{1, \ldots, k\} \). Since there are two possible choices of \( c_i \in \{0, 1\} \) for each \( i \), the number of \( \ell \)-level subfunctions of \( f_n \circ g_n \) is at most \( 2^k \).

Case 2: \( \ell = 2k - 1 \) for \( k \in \{1, \ldots, n\} \). There exist \( a_1, b_1, \ldots, a_{k-1}, b_{k-1}, a_k \in \{0, 1\} \) such that

\[
h(y_k, x_{k+1}, y_{k+1}, \ldots, z_m) = (f_n \circ g_n)(a_1, b_1, \ldots, a_k, y_k, x_{k+1}, y_{k+1}, \ldots, z_m)
\]

\[
= \begin{cases} 
  c_i & \text{if } \beta(z_1, \ldots, z_m) = i < k; \\
  a_k \circ y_k & \text{if } \beta(z_1, \ldots, z_m) = k; \\
  x_i \circ y_i & \text{if } \beta(z_1, \ldots, z_m) = i > k,
\end{cases}
\]

where \( c_i = a_i \circ b_i \) for \( i \in \{1, \ldots, k-1\} \). Since there are two possible choices of both \( c_i \in \{0, 1\} \) for each \( i \) and \( a_k \), the number of \( \ell \)-level subfunctions of \( f_n \circ g_n \) is at most \( 2^k \).

Case 3: \( \ell = 2n + k \) for \( k \in \{1, \ldots, m\} \). In this case, \( h \) is a function of \( m - k \) variables. Therefore, the number of \( \ell \)-level subfunctions of \( f_n \circ g_n \) is at most \( 2^{2n-k} \) by Lemma 3.

All in all, \( f_n \circ g_n \) has at most

\[
\sum_{k=0}^{n} 2^k + \sum_{k=1}^{n} 2^k + \sum_{k=1}^{m} 2^{2n-k} \leq 2 \cdot 2^n + 2 \cdot 2^n + 2 \cdot 2^n = 6 \cdot 2^n
\]

subfunctions. \( \square \)

Corollary 8. The number of nodes of \( B(f_n \circ g_n) \) is at most \( 6 \cdot 2^n \) for any \( \circ \).

Figure 1 illustrates the linear growth in the BDD size through computation of \( \text{Apply}(\circ, B(f), B(g)) \).

Theorem 9. For any properly binary Boolean operation \( \circ \), Algorithm 1 does not run in \( O(|B(f_n)| + |B(g_n)|) + |B(f_n \circ g_n)|) \) time on input \( \langle \circ, B(f_n), B(g_n) \rangle \).

Proof. By Lemmata 4 and 6 and Corollary 8. \( \square \)
3.2. More Drastic Counterexample

The size of the output BDD $B(f_n \circ g_n)$ is proportional to the input size $|B(f_n)| + |B(g_n)|$ in the construction in the previous subsection. The construction can be modified so that the resultant function is constant when the binary operation $\circ$ is $\land$ or $\lor$. For a positive integer $n$, we assume that we have $2n + m + 1$ variables $x_1, y_1, x_2, y_2, \ldots, x_n, y_n, z_1, \ldots, z_m, w$ in this order, where $m = \lceil \log n \rceil$. We then define two functions of these variables $f_n^i$ and $g_n^i$ for $i = 0, 1$ by

$$f_n^i(x_1, y_1, \ldots, z_m, w) = \begin{cases} f^i_n(x_1, y_1, \ldots, z_m) & \text{if } w = 0; \\ i & \text{if } w = 1; \end{cases}$$

$$g_n^i(x_1, y_1, \ldots, z_m, w) = \begin{cases} i & \text{if } w = 0; \\ g_n(x_1, y_1, \ldots, z_m) & \text{if } w = 1, \end{cases}$$

where $f_n$ and $g_n$ are the functions given in Sec. 3.1. Clearly,

$$|B(f_n^i \land g_n^i)| = |B(f_n^i \lor g_n^i)| = 1,$$

because

$$(f_n^0 \land g_n^0)(x_1, y_1, \ldots, z_m, w) = 0,$$

$$(f_n^1 \lor g_n^1)(x_1, y_1, \ldots, z_m, w) = 1,$$

for any $x_1, y_1, \ldots, x_n, y_n, z_1, \ldots, z_m, w$. All lemmata in Sec. 3.1 still hold for the computation of $B(f_n^i \land g_n^i)$ and $B(f_n^i \lor g_n^i)$. Specifically, the numbers of nodes of $B(f_n^0)$ and $B(g_n^0)$ are at most $4 \cdot 2^n$ (Lemma 4)\(^2\) and Algorithm 1 recursively calls itself at least

\[^3\text{In fact, the shape of } B(f_n^0) \text{ is almost identical to that of } B(f_n). \text{ We can obtain } B(f_n^0) \text{ from } B(f_n) \text{ by reconnecting all edges going into terminal node 1 to a new node assigned } \langle w, 1, 0 \rangle. \text{ That is, } |B(f_n^0)| = |B(f_n)| + 1. \text{ The same argument holds for } f_n^1, g_n^1, g_n^0.\]

---

Figure 1: The result of $\text{Apply}(\oplus, p, q)$, where $p$ is labeled by $x_i$ and $q$ by $y_i$. Here 0-edges are shown by dotted lines and 1-edges are by solid lines.
Table 1: Sizes of $B(f_n)$, $B(g_n)$, $B(f_n \land g_n)$, and $B(f_n \lor g_n)$, and computation times of $f_n \land g_n$ and $f_n \lor g_n$. $T^c_n$ ($T^s_n$) is the computation time (sec.) of $f_n \land g_n$ by the SAPPORO BDD (CUDD) package.

| $n$ | $|B(f_n)|$ | $|B(g_n)|$ | $|B(f_n \land g_n)|$ | $T^c_n$ | $T^s_n$ | $|B(f_n \lor g_n)|$ | $T^c_n$ | $T^s_n$ |
|-----|-----------|-----------|---------------------|--------|--------|---------------------|--------|--------|
| 9   | 895       | 895       | 1.661               | 0.031  | 0.046  | 1.406               | 0.04   | 0.04   |
| 10  | 1,662     | 1,662     | 3.196               | 0.141  | 0.187  | 2.685               | 0.156  | 0.156  |
| 11  | 3,196     | 3,196     | 6.266               | 0.656  | 0.780  | 5.243               | 0.624  | 0.686  |
| 12  | 6,264     | 6,264     | 12.406              | 2.81   | 3.23   | 10.359              | 2.50   | 2.78   |
| 13  | 12,400    | 12,400    | 24,686              | 12.3   | 13.3   | 20,591              | 11.1   | 11.5   |
| 14  | 24,672    | 24,672    | 49,246              | 54.5   | 53.5   | 41,055              | 49.1   | 46.0   |
| 15  | 49,216    | 49,216    | 98,366              | 243    | 222    | 81,983              | 227    | 196    |
| 16  | 98,304    | 98,304    | 196,606             | 1038   | 1043   | 163,839             | 1068   | 903    |

$(2^n)^2$ times for inputs $\langle \land, B(f_n), B(g_n) \rangle$ and $\langle \lor, B(f_n), B(g_n) \rangle$ (Lemma 6).

4. Experimental Results on Actual BDD Packages

Many existing implementations adopt slight modifications to the original BDDs introduced in Sec. 2, which make theoretical analysis on the time complexity cumbersome, although those modifications are usually supposed to have no significant change on the computational complexity of BDDs. This section demonstrates experimentally that our counterexamples are most likely valid for existing packages. To this end, using the SAPPORO BDD package by Minato (unreleased) and the CUDD package by Somenzi [5], we computed $f_n \land g_n$, $f_n \lor g_n$, and $f_0 \land g_0$, where $f_n$, $g_n$, $f_0$, and $g_0$ are defined in Sec. 3. Tables 1 and 2 show the sizes of the BDDs and the computation time for these binary operations when $n = 9, \ldots, 16$. The program was coded in C++ and was performed on a 2.80 GHz CPU with 6 GB RAM, running Windows 7 with Cygwin. (The source code is available from http://www-erato.ist.hokudai.ac.jp/~jkawahara/BryantConjecture.html.) For each increase of $n$ by one, $|B(f_n)|$, $|B(g_n)|$, $|B(f_n \land g_n)|$ and $|B(f_n \lor g_n)|$ increase by a factor of about two, while the computation time increases by a factor of about four. Thus, the results suggest that the computation time of these binary operations are not bound linearly by the size of input and output BDDs.

5. Concluding Remarks

In this article, we have presented counterexamples to the long-standing conjecture on the complexity of the BDD $\text{Apply}$ algorithm. Our results show that the computation time of the algorithm is not always a linear relation of the input-output sizes for any properly binary operation ($f \circ g$). Our counterexamples require a computation time that is a quadratic function of the input BDD size, while the output BDD size is linearly related to input size. In addition, for the operations $\land$ and $\lor$, we have shown a stronger counterexample where the output BDD size is constant while computation time is still a quadratic of the input BDD size.

We note that our results can also be applied to the zero-suppressed binary decision diagram (ZDD) [3], which is based on node reduction rules that are slightly different.
Table 2: Sizes of $B(f^0_n)$, $B(g^0_n)$, and $B(f^0_n \land g^0_n)$, and computation times of $f^0_n \land g^0_n$. $T^S$ ($T^C$) is the computation time (s) of $f^0_n \land g^0_n$ by the SAPPORO BDD (CUDD) package.

| $n$ | $|B(f^0_n)|$ | $|B(g^0_n)|$ | $|B(f^0_n \land g^0_n)|$ | $T^S$ | $T^C$ |
|-----|--------------|--------------|--------------------------|------|------|
| 9   | 1,278        | 1,278        | 1                        | 0.032| 0.016|
| 10  | 2,300        | 2,300        | 1                        | 0.094| 0.078|
| 11  | 4,344        | 4,344        | 1                        | 0.390| 0.296|
| 12  | 8,432        | 8,432        | 1                        | 1.50 | 1.14 |
| 13  | 16,608       | 16,608       | 1                        | 6.77 | 4.66 |
| 14  | 32,960       | 32,960       | 1                        | 29.7 | 20.7 |
| 15  | 65,664       | 65,664       | 1                        | 125  | 87.8 |
| 16  | 131,072      | 131,072      | 1                        | 621  | 393  |

from the ones in conventional BDD. Space does not allow to discuss ZDD here; however, we consider that the difference between ZDD and BDD has almost no (exponentially small) effect on the computation time of our counterexamples. Thus, asymptotic behaviors will be the same in our argument for ZDD.

While our results are important from a theoretical point of view, the following observations can also be made.

- Our counterexamples have an artificial structure, and would rarely appear in real-life problems. In most practical situations, the conjecture may be correct.
- While our counterexamples assume a specific variable ordering for the Boolean functions, the ordering is far from optimal. If we use a reasonable variable ordering, the conjecture may still stand. We are yet to find a counterexample applicable regardless to the variable ordering used.

Although theoretically disproved, the conjecture is still useful as a good estimation of the computation time for BDD construction in many practical applications.

References