On the BDD of a Random Boolean Function

Jean Vuillemin Fredéric Béal.*

October 13, 2003

Abstract

The *Binary Decision Diagram* BDD, the *Binary Moment Diagram* BMD and the *Minimal Deterministic Automaton* MDA are three canonical representations for Boolean functions. Exact expression $a(i)$ and $w(i)$ are provided for the average and worst size BDD over $\mathbf{B}^i \mapsto \mathbf{B}$, and they are proved equal to those for the average and worst size BMD. The expressions $\overline{a}(i)$ and $\overline{w}(i)$ for MDA are just slightly bigger since

$$1 = \lim_{i \mapsto \infty} \frac{\overline{a}(i)}{a(i)} = \lim_{i \mapsto \infty} \frac{\overline{w}(i)}{w(i)}.$$

The significant differences between worst and average sizes are shown located on levels c and $c+1$: the *critical depth* $c = \lfloor r \rfloor$ is the integer part of the root $r \in \mathbf{R}$ of $i = r + 2^r$. The analysis shows that the average to worst size ratios $\frac{a(i)}{w(i)}$ and $\frac{\overline{a}(i)}{\overline{w}(i)}$ oscillate between

$$1 - \frac{1}{2e} = 0.81606\cdots \text{ and } 1 \text{ as } i \mapsto \infty.$$

Successive minima are found for $i = n + 2^n$ and $n \in \mathbf{N}$, where the gain between average and worst sizes is about 18%. Yet, such numbers are far in between: the gain far less than 1%, with probability one over the integers. The BDD/BMD/MDA sizes of a random function with a random number i of inputs are all equivalent to those of the worst size structure $a(i) \simeq \overline{a}(i) \simeq w(i) \simeq \overline{w}(i)$.

1 Introduction

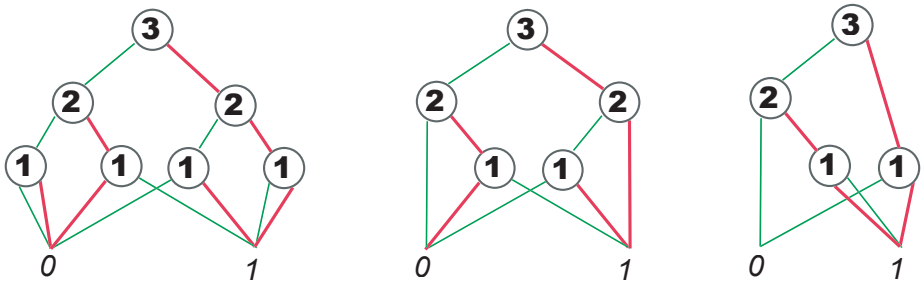
The MDA, BDD and BMD are classical [1, 2, 3] *strong normal forms* for representing Boolean functions, and they are key to modern circuit verification.

This paper provides an exact analysis for the worst and average size of these three structures. The worst case analysis relates to counting arguments which are well known for circuits [6]. Our contribution lies in the exact and asymptotic analysis of the average sizes.[1]

*Ecole Normale Supérieure, 45 rue d'Ulm, 75005 Paris France.

[1] Bryant's publication [1] on BDDs is reputed to be the most often quoted in Computer Science. Thus, we expect that the present analysis of the random BDD may have already been published, for whole or in part. We will happily retract/amend/acknowledge our claims, based on all prior-art provided by our kind readers.

The example function is defined by the expression $E = x'_1x_2 + x_1x_3$, with $x'_k = 1 - x_k$.



Graphs of $MDA_3(E)$, $BDD_3(E)$ and $BMD_3(E)$.

$$\begin{aligned} BDD_3(E) &= \{0, 1, x_1, x'_1, x_2x'_1, x_1 \cup x_2, E\}, \\ BMD_3(E) &= \{0, 1, x_1, x'_1, x_2x'_1, E\}. \end{aligned}$$

Figure 1: Example MDA, BDD and BMD.

2 Boolean Functions and Canonical Forms

Let $\mathbf{B}_i = \mathbf{B}^i \mapsto \mathbf{B}$ denote the set formed by the $\beta_i = |\mathbf{B}_i| = 2^{2^i}$ Boolean functions having $i \in \mathbf{N}$ inputs and a single output. An *abstract* Boolean function $f^e = [\![e]\!] \in \mathbf{B}_i$ is represented by some *concrete* expression (a.k.a. circuit net-list)

$$e \in \mathbf{E}_i = \langle 0, 1, x_1, \cdots, x_i, \neg, \cap, \cup, \oplus \rangle$$

composed from 0, 1, inputs $\overline{x} = x_1 \cdots x_i$ and say, the logical *not, and, or, xor*. In turn, each expression $e \in \mathbf{E}_i$ denotes a unique function $f^e(\overline{x}) = [\![e]\!] \in \mathbf{B}_i$. The Boolean value $f^e(\overline{b}) = [\![e_{\overline{x}=\overline{b}}]\!] \in \mathbf{B} = \{0, 1\}$ is computed by substituting $\overline{b} = b_1 \cdots b_i \in \mathbf{B}^i$ for $\overline{x} = x_1 \cdots x_i$ in e and by simplifying, to either 0 or 1.

Let $e \in \mathbf{E}_i$ be some Boolean expression with i inputs, and $f^e = [\![e]\!] \in \mathbf{B}_i$ be the denoted Boolean function. For $i > 0$ and $b \in \mathbf{B}$, the partial substitutions $e_{x_i=b}$ project f^e on two *prefix functions* with $i - 1$ inputs:

$$f^e_{x_i=b} = [\![e_{x_i=b}]\!] \in \mathbf{B}_{i-1}.$$

A function $f \in \mathbf{B}_i$ is thus uniquely expressed by Shannon's prefix decomposition:

$$f = x_i f_{x_i=1} + (1 - x_i) f_{x_i=0}.$$

The MDA for $f \in \mathbf{B}_i$ can be constructed as follows:

1. Recursively apply Shannon's prefix decomposition down to the constant functions $0, 1 \in \mathbf{B}_0$; the result is a complete binary decision tree of depth i, whose 2^i leaves are labelled by the 2^i bits: $f(\overline{b}) \in \mathbf{B}$ for $\overline{b} \in \mathbf{B}^i$.

2. Systematically share the nodes which represent equal Boolean functions, at all levels p for $1 \leq p \leq i$ in the decision tree.

The resulting MDA data structure is a *directed acyclic graph* in which: all paths from the root to a leaf have the same length i; distinct nodes $e \neq e' \in \mathbf{E}_p$ at depth p represent distinct Boolean functions: $f^e(\overline{b}) = 1 - f^{e'}(\overline{b})$ for some $\overline{b} \in \mathbf{B}^p$.

The *prefix closure* of $f \in \mathbf{B}_i$ is the set $MDA_i(f)$ of Boolean functions which label nodes in the MDA for f. It is recursively defined by:

$$\begin{array}{rcl} MDA_0(f) & = & \{f\} \text{ for } f = 0 \text{ and } f = 1, \\ i > 0: \ MDA_i(f) & = & \{f\} \cup MDA_{i-1}(f_{x_i=0}) \cup MDA_{i-1}(f_{x_i=1}). \end{array}$$

The partial derivative of $f \in \mathbf{B}_i$ with respect to variable x_i is defined by

$$\frac{\partial f}{\partial x_i} = (f_{x_i=0} \oplus f_{x_i=1}), \text{ so that } (\frac{\partial f}{\partial x_i} = 0) \Leftrightarrow (f = f_{x_i=0} = f_{x_i=1}).$$

Condition $\frac{\partial g}{\partial x_p} = 0$ detects when g is independent of input bit x_p. We let $\mathbf{B}'_p = \{g \in \mathbf{B}_p : \ \frac{\partial g}{\partial x_p} \neq 0\}$ denote the $\beta'_p = |\mathbf{B}'_p| = \beta_p - \beta_{p-1}$ functions in $\mathbf{B}_p$ which effectively depend upon input bit x_p.

The BDD for $f \in \mathbf{B}_i$ may be constructed from the MDA, by simplifying away all nodes $g \notin \mathbf{B}'_p$ with are independent of x_p. It is recursively defined by:

$$\begin{array}{rcl} BDD_0(f) & = & \{f\}, \\ \frac{\partial f}{\partial x_i} = 0: \ BDD_i(f) & = & BDD_{i-1}(f_{x_i=0}), \\ \frac{\partial f}{\partial x_i} \neq 0: \ BDD_i(f) & = & \{f\} \cup BDD_{i-1}(f_{x_i=0}) \cup BDD_{i-1}(f_{x_i=1}). \end{array}$$

A dual of Shannon's decomposition is the Reed-Muller decomposition:

$$f = f_{x_i=0} \oplus x_i \frac{\partial f}{\partial x_i}.$$

The BMD for $f \in \mathbf{B}_i$ is constructed by recursively applying Reed-Muller's decomposition and by systematically sharing all common sub-expressions:

$$\begin{array}{rcl} BMD_0(f) & = & \{f\}, \\ \frac{\partial f}{\partial x_i} = 0: \ BMD_i(f) & = & BMD_{i-1}(f_{x_i=0}), \\ \frac{\partial f}{\partial x_i} \neq 0: \ BMD_i(f) & = & \{f\} \cup BMD_{i-1}(f_{x_i=0}) \cup BMD_{i-1}(\frac{\partial f}{\partial x_i}). \end{array}$$

3 Worst Case Analysis

In defining the size of our structures, it is convenient to only count internal nodes and to exclude the two leaf nodes $0, 1 \in \mathbf{B}_0$, as in the tables in Figure 2.

Definition 1 *Let the worst size MDA, BDD, BMD over $f \in \mathbf{B}_i$ be:*

$$\begin{array}{rcl} W^{mda}(i) & = & \max\{|MDA_i(f)| : \ f \in \mathbf{B}_i\}, \\ W^{bdd}(i) & = & \max\{|BDD_i(f)| : \ f \in \mathbf{B}_i\}, \\ W^{bmd}(i) & = & \max\{|BMD_i(f)| : \ f \in \mathbf{B}_i\}. \end{array}$$

$i\backslash p$	1	2	3	4	$w(i)$
1	1				1
2	2	1			3
3	2	2	1		5
4	2	4	2	1	9
5	2	8	4	2	17
6	2	12	8	4	29
7	2	12	16	8	45
8	2	12	32	16	77

Table for $w_p(i)$ and $w(i)$.

$i\backslash p$	1	2	3	4	$\overline{w}(i)$
1	1				1
2	2	1			3
3	4	2	1		7
4	4	4	2	1	11
5	4	8	4	2	19
6	4	16	8	4	35
7	4	16	16	8	51
8	4	16	32	16	83

Table for $\overline{w}_p(i)$ and $\overline{w}(i)$.

Figure 2: Worst Size tables.

The analysis for $i \in \mathbf{N}$ relates to the unique root $r = r(i) \in \mathbf{R} \geq 0$ of[2]

$$i = r + 2^r. \tag{1}$$

The *critical depth* $c = c(i) \in \mathbf{N}$ of i is the integer part $c = \lfloor r \rfloor$ of r.

3.1 Exact Analysis

In the worst case structures of depth $i \in \mathbf{N}$, the nodes above $c+1$ form a complete binary tree. The nodes beneath c enumerate all the Boolean functions in $\mathbf{B}_c$ - in a redundant way within the MDA, non-redundant within the BDD/BMD.

Proposition 1 *The worst BDD and BMD over $f \in \mathbf{B}_i$ have equal sizes*

$$w(i) = W^{bdd}(i) = W^{bmd}(i) = 2^{i-c} + 2^{2^c} - 3 \tag{2}$$

and $c = c(i)$ is the critical depth of i. The worst MDA has the related size

$$\overline{w}(i) = W^{mda}(i) = w(i) + \beta''_{c-1} \tag{3}$$

for $\beta''_j = \sum_{0<p\leq j} \beta_p$.

Proof: Let $\overline{w}_p(i)$ count the nodes at depth p in the worst size MDA_i, so that $\overline{w}(i) = \sum_p \overline{w}_p(i)$. There are 2^{i-p} nodes at depth p in a complete binary decision tree of depth i. In the worst case, each node represents a different Boolean function and $\overline{w}_p(i) = 2^{i-p}$. This is true for as long as there are enough functions to choose from, namely $2^{i-p} \leq \beta_p$. Otherwise, the worst case $\overline{w}_p(i) = \beta_p$ takes place when each Boolean function in $\mathbf{B}_p$ is represented by some node at depth p in MDA_i. In summary

$$\overline{w}_p(i) \quad = \quad \min(2^{i-p}, \beta_p).$$

[2]The root of (1) is related to that $x = L(y)$ of Lambert's transcendental equation $xe^x = y$ by $r = i - \frac{2^i}{e^{L(\ln(2)2^i)}}$ - function L is called *LambertW* by [5].

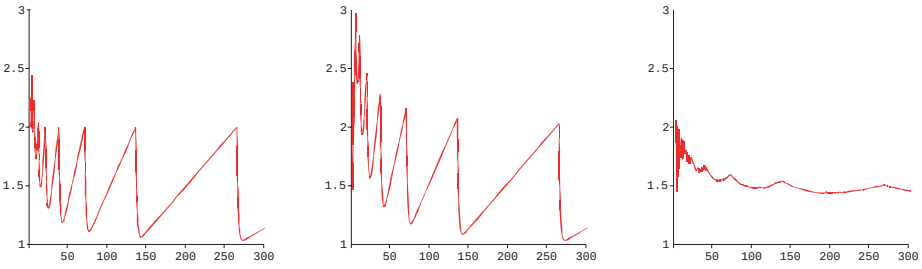


Figure 3: Graphs of $\frac{w(i)}{\beta_{r(i)}}$, $\frac{iw(i)}{2^i}$ and $\frac{1}{i}\sum_{1\leq n\leq i}\frac{nw(n)}{2^n}$.

The number of nodes at depth p in the worst BDD_i or BMD_i is then

$$w_p(i) \quad = \quad \min(2^{i-p}, \beta'_p),$$

since $\beta'_p = \beta_p - \beta_{p-1}$ counts there the Boolean functions in $\mathbf{B}'_p$.

The sign of $2^{i-p} - \beta_p$ is equal to the sign of $d_p = i - p - 2^p$ and to the sign of $p - r$ since $d_p = p - r + 2^{p-r}$. The following equivalence

$$(2^{i-p} < \beta_p) \Leftrightarrow (d_p > 0) \Leftrightarrow (p > r) \Leftrightarrow (2^{i-p} < \beta'_p)$$

is simply derived for $p > 0$. Substituting in the above expressions gives us:

$$\begin{aligned} \overline{w}_p(i) \quad &= \quad \begin{cases} \beta_p \text{ if } p \leq r, \\ 2^{i-p} \text{ if } p > r, \end{cases} \\ \text{and } w_p(i) \quad &= \quad \begin{cases} \beta'_p \text{ if } p \leq r, \\ 2^{i-p} \text{ if } p > r. \end{cases} \end{aligned}$$

Summing up $w(i) = \sum w_p(i)$ and $\overline{w}(i) = \sum \overline{w}_p(i)$ yield (2,3). *Q.E.D.*

3.2 Asymptotic Analysis

One classical [6] asymptotic equivalent to $w(i)$ is $\frac{2^i}{i}$; yet Figure 3 indicates that the ratio $\frac{iw(i)}{2^i}$ has no limit for $i \mapsto \infty$.

Proposition 2 *For i large enough, the size $\overline{w}(i)$ of worst MDA is equivalent to that $w(i)$ of worst BDD/BMD:* $1 = \lim_{i\mapsto\infty} \frac{\overline{w}(i)}{w(i)}$.

The ratios between worst-size and $\frac{2^i}{i}$ (or $\beta_r = 2^{2^r}$ for $r + 2^r = i$) oscillate:

$$1 = \liminf_{i\mapsto\infty} \frac{w(i)}{\beta_r} \qquad \textit{and likewise for } \frac{iw(i)}{2^i}, \frac{\overline{w}(i)}{\beta_r} \textit{ and } \frac{i\overline{w}(i)}{2^i};$$

$$2 = \limsup_{i\mapsto\infty} \frac{w(i)}{\beta_r} \qquad \textit{and likewise for the above ratios.}$$

The average ratio is half way between the limiting values:

$$\frac{3}{2} = \lim_{i\mapsto\infty} \frac{1}{i} \sum_{1\leq n\leq i} \frac{w(n)}{\beta_{r(n)}} \qquad \textit{and likewise for the above ratios.}$$

Proof: Expression (3) gives the equivalence

$$\frac{\overline{w}(i)}{w(i)} = 1 + \frac{\beta''_{c-1}}{w(i)} = 1 + O(\frac{1}{\beta_{c-1}}) = 1 + O(\frac{1}{\sqrt{w(i)}})$$

whose limit is sharply 1 for $i, c \mapsto \infty$.

Let $f = r - c$ be the fractional part of $r = r(i)$. From (2), we derive that $w(i) + 3 = 2^{i-c} + \beta_c = \beta_r(2^f + \beta_c^{1-2^f}) = \beta_r g(2^f - 1, 1/\beta_c)$ where

$$g(t, x) = 1 + t + x^t.$$

Function g decreases from $g(0, x) = 2$ to $g_{min} = g(t_x, x)$ as t increases from 0 to t_x; g then increases from g_{min} to its maximum $g(1, x) = 2 + x$ as t goes from t_m to 1. The minimum is reached at $x^{t_x} \log \frac{1}{x} = 1$ and $g_{min} = 1 + O\left(\frac{\log\log \frac{1}{x}}{\log \frac{1}{x}}\right)$. Since $\lim_{x \mapsto 0} g(1, x) = 2$ and $\lim_{x \mapsto 0} g(t_x, x) = 1$, we conclude that

$$1 = \liminf_{x_n \mapsto 0} g(t_n, x_n) \quad \text{and} \quad 2 = \limsup_{x_n \mapsto 0} g(t_n, x_n)$$

over all real sequences $t_n \in [0, 1]$, hence the claimed limits for $\frac{w(i)}{\beta_r} = g(t, x)$.

Since $\overline{w}(i) \simeq w(i)$, the same limits apply to $\frac{\overline{w}(i)}{\beta_r}$ and to $\frac{i\overline{w}(i)}{2^i}$ as

$$\beta_r = 2^{2^r} = \frac{2^i}{i - r} = \frac{2^i}{i}(1 + O(\frac{\log(i)}{i})).$$

Similarly for $E(n) \in \{\frac{w(n)}{\beta_{r(n)}}, \frac{nw(n)}{2^n}, \frac{\overline{w}(n)}{\beta_{r(n)}}, \frac{n\overline{w}(n)}{2^n}\}$, all $\lim_{i \mapsto \infty} \frac{1}{i} \sum_{1 \le n \le i} E(n)$ have a limit equal to L. The value $L = \frac{3}{2}$ is obtained for $c \mapsto \infty$, from the finite sum over integers which have critical depth c:

$$\frac{1}{1 + 2^c} \sum_{c(n)=c} \frac{nw(n)}{2^n} = \frac{1}{1 + 2^c} \sum_{d \le 2^c} (1 + (c + d)2^{-c})(1 + 2^{-d}) = 1 + \frac{1}{2} + O(1/c).$$

The analysis confirms the intuition from Figure 3: in the limit, the ratio $\frac{iw(i)}{2^i}$ tends to the piece-wise-linear pseudo-periodic function $\rho_\infty(i) = \frac{i}{c + 2^c}$, where $c \in \mathbf{N}$ is the critical integer such that $C(c) = c + 2^c \le i < C(c + 1)$. *Q.E.D.*

4 Average Case Analysis

Definition 2 *Let the average size MDA, BDD, BMD over $f \in \mathbf{B}_i$ be:*

$$\begin{aligned} A^{mda}(i) &= \frac{1}{\beta_i} \sum_{f \in \mathbf{B}_i} |MDA_i(f)|, \\ A^{bdd}(i) &= \frac{1}{\beta_i} \sum_{f \in \mathbf{B}_i} |BDD_i(f)|, \\ A^{bmd}(i) &= \frac{1}{\beta_i} \sum_{f \in \mathbf{B}_i} |BMD_i(f)|. \end{aligned}$$

$q\backslash p$	1	2	3	$4\cdots$	$oa(q)$
1	1				1
2	1.7	1			2.75
3	2.7	1.9	1		5.7
4	3.6	3.6	2	1	10.2
5	4.0	6.5	4.0	2.0	17.4
6	4.0	10.3	7.9	4.0	29.2
7	4.0	14.0	15.5	8.0	48.5
8	4.0	15.7	30.1	16.0	80.9

Table for $\overline{a}_p(q)$ and $\overline{a}(q)$.

$q\backslash p$	1	2	3	$4\cdots$	$a(q)$
1	0.5				0.5
2	0.9	0.8			1.6
3	1.4	1.5	0.9		3.8
4	1.8	2.7	1.9	1.0	7.4
5	2.0	4.8	3.7	2.0	13.5
6	2.0	7.7	7.4	4.0	24.1
7	2.0	10.5	14.6	8.0	42.
8	2.0	11.8	28.3	16.0	73.

Table for $a_p(q)$ and $a(q)$.

Figure 4: Average size tables.

Let $A_p^{mda}(i)$ count the nodes at depth p in $A^{mda}(i) = \sum_p A_p^{mda}(i)$, and similarly for BDD and BMD.

The average analysis is related to that of *hashing* [4]. Let $\mathcal{N} = [n_1 \cdots n_k]$ be some sequence of k integers chosen at random in $\{0 \cdots m-1\}$. The probability that an integer j such that $0 \leq j < m$ belongs to $\mathcal{N}$ is

$$\Pr(j \in \mathcal{N}) = 1 - (1 - \frac{1}{m})^k = h(\frac{1}{m}, k). \tag{4}$$

The *hash* function to analyze here is $h(x, y) = 1 - (1 - x)^y$.

4.1 Exact Analysis

Proposition 3 *The average BDD and BMD have the same size at depth p:*

$$a_p(i) = A_p^{bdd}(i) = A_p^{bmd}(i) = \beta'_p h(x_p, y_p) \tag{5}$$

for all $1 \leq p \leq i$; here, $x_p = \frac{1}{\beta_p}$ and $y_p = 2^{i-p}$. The size of the average MDA is

$$\overline{a}_p(i) = A_p^{mda}(i) = \beta_p h(x_p, y_p). \tag{6}$$

Proof: The probability that some Boolean function $g \in \mathbf{B}_p$ is $g \in MDA_i(f)$ among the 2^{i-p} prefixes of a random $f \in \mathbf{B}_i$ amounts to $h(x_p, y_p)$ by (4). The average number of nodes at depth p in $MDA_i(f)$ is the sum $\overline{a}_p(i) = \beta_p h(x_p, y_p)$ of these probabilities over $\mathbf{B}_p$. Summing over $\mathbf{B}'_p$ yields $a_p(i) = \beta'_p h(x_p, y_p)$ for $BDD_i(f)$ and as well for $BMD_i(f)$. *Q.E.D.*

4.2 Asymptotic Analysis

The limit (if any) of the hash function $h(x, y) = 1 - (1 - x)^y$ for $x \mapsto 0$ and $y \mapsto \infty$ depends on that (if any) of the product $p = xy$: $h(x, y) \mapsto 0$ if $p \mapsto 0$; $h(x, y) \mapsto 1$ if $p \mapsto \infty$; finally $h(x, y) \mapsto 1 - \frac{1}{e^p}$ if $xy \mapsto p \in \mathbf{R} > 0$.

Lemma 1 *Let $i \in \mathbf{N}$ have critical depth $c = \lfloor r(i) \rfloor$; let $x_p = \frac{1}{\beta_p} < 1$ and $y_p = 2^{i-p} \geq 1$. The number $h(x_p, y_p) = 1 - (1 - x_p)^{y_p} = 1 - e^{y_p \log(1-x_p)}$ equals*

$$h(x_p, y_p) = \begin{cases} 1 - \theta x_{c+1} & \text{if } 1 \leq p < c; \\ 1 - e^{-x_p y_p} - \theta x_p & \text{if } p = c; \\ x_p y_p (1 - \frac{x_p y_p}{2}(1 - \frac{\theta}{6})) & \text{if } p = c+1; \\ x_p (y_p - \theta) & \text{if } c+1 < p \leq i. \end{cases}$$

Proof: Throughout this paper, the reader should see θ_n whenever *she/he/it* reads the letter θ. The *invisible index* n is the number of occurrences of θ before this point in the text. Each variable θ_n represents a real number such that $0 < \theta_n < 1$, and no relation is assumed beyond that. The usage of θ is restricted to a single occurrence per real-valued expression.

By (1), the sign of $d_p = i - p - 2^p$ is equal to the sign of $r - p$ since

$$d_p = \log_2(x_p y_p) = (r - p) + (2^r - 2^p).$$

- The condition $p < c \Leftrightarrow p \leq r - 1 \Leftrightarrow d_p > 2^p$ implies that

  $$1 - e^{y_p \log(1-x_p)} = 1 - e^{-x_p y_p/(1+\theta)} = 1 - \theta e^{-x_p y_p/2} = 1 - \theta x_{c+1};$$

  indeed, $x_p y_p/2 = 2^{d_p - 1} > 2^{c+1}$ follows from $d_p - 1 \geq 2^{c-1} > c+1$ which is true for $c > 2$. A computer verification confirms the expression for $c \leq 2$.

- The condition $p = c \Leftrightarrow r - 1 < p \leq r \Leftrightarrow 0 \leq d_p \leq 2^c$ implies that

  $$e^{y_c \log(1-x_c)} = e^{-x_c y_c + \frac{\theta}{2} x_c^2 y_c} = e^{-x_c y_c}(1 + \theta x_c^2 y_c) = e^{-x_c y_c} + \theta x_c$$

  since $\frac{\theta}{2} x_c^2 y_c = \frac{\theta}{2} 2^{d_c - 2^c} < 1/2$ and $x_c y_c < e^{x_c y_c}$.

- The condition $p > c \Leftrightarrow p > r \Leftrightarrow d_p < 0 \Leftrightarrow x_p y_p < 1$ implies that

  $$1 - e^{-x_p y_p + \theta x_p} = 1 - e^{-x_p y_p} - \theta x_p = x_p y_p (1 - \frac{x_p y_p}{2}(1 - \frac{\theta}{6})).$$

  Condition $p > c + 1 \Leftrightarrow p > r + 1 \Leftrightarrow d_p + 1 < -2^{p-1}$ finally entails

  $$h(x_p, y_p) = x_p y_p - \theta (x_p y_p)^2 = x_p y_p - \theta x_p$$

  since $x_p y_p < x_{p-1}$ implies that $(x_p y_p)^2 < x_{p-1}^2 = x_p$. *Q.E.D.*

4.3 Average versus Worst Sizes

The tables in Figures 2 and 4 indicate that the average and worst cases have almost equal sizes, except near the critical depth $c = c(i)$. Indeed, Lemma 1 implies that the differences $w_p(i) - a_p(i)$ and $\overline{w}_p(i) - \overline{a}_p(i)$ are infinitesimals unless $p = c$ or $p = c + 1$. In other words, the average size structure is the same as the worst size structure at all levels, except sometimes at depths c or $c + 1$.

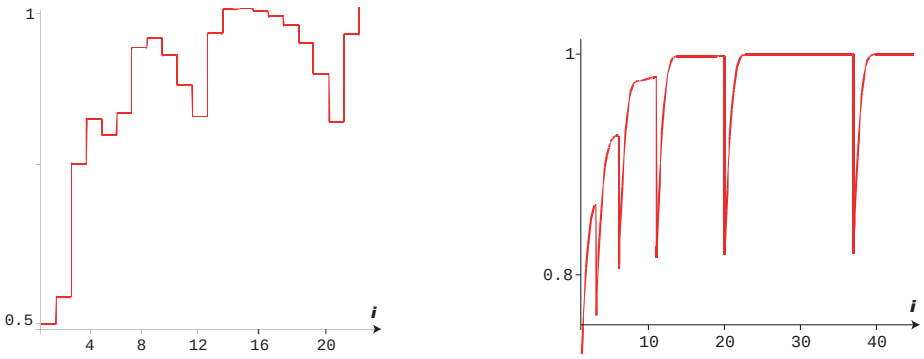


Figure 5: Graphs of $\rho(i) = \frac{a(i)}{w(i)}$ and $\overline{\rho}(i) = \frac{\overline{a}(i)}{\overline{w}(i)}$.

The difference between worst and average size is maximized when the number i of inputs is of the critical form $i = n + 2^n$, for $n \in \mathbf{N}$. In this critical case, the ratio between worst and average size quickly approaches $1 - \frac{1}{2e} = 0.8160\cdots$ for a maximal gain of 18%. Yet, it follows from (10) that this ratio is greater than 0.999 for *almost all* $i \in \mathbf{N}$: on average, the gain is thus far less than 1%.

Proposition 4 *For i large enough, the average BDD, BMD and MDA all have the same relative size:*

$$1 \quad = \quad \lim_{i \mapsto \infty} \frac{\overline{a}(i)}{a(i)}, \tag{7}$$

The average to worst size ratios $\rho(i) = \frac{a(i)}{w(i)}$ and $\overline{\rho}(i) = \frac{\overline{a}(i)}{\overline{w}(i)}$ are such that:

$$1 \quad = \quad \limsup_{i \mapsto \infty} \rho(i) = \limsup_{i \mapsto \infty} \overline{\rho}(i), \tag{8}$$

$$1 - \frac{1}{2e} \quad = \quad \liminf_{i \mapsto \infty} \rho(i) = \liminf_{i \mapsto \infty} \overline{\rho}(i), \tag{9}$$

$$1 \quad = \quad \lim_{i \mapsto \infty} \frac{1}{i} \sum_{1 \le n \le i} \rho(n) = \lim_{i \mapsto \infty} \frac{1}{i} \sum_{1 \le n \le i} \overline{\rho}(n). \tag{10}$$

Proof: Let $c = c(i)$ be the crital depth, so that $i = c + d + 2^c$ and $0 \le d \le 2^c$. We split the sum $\overline{a}(i) = \sum_{p<c} \overline{a}_p(i) + \overline{a}_c(i) + \overline{a}_{c+1}(i) \sum_{p>c+1} \overline{a}_p(i)$ and replace each term $\overline{a}_p(i) = \beta_p h(x_p, y_p)$ by its equivalent from Lemma 1 to find that

$$\overline{a}(i) \quad = \quad \overline{w}(i) - \overline{\delta}(i) - \theta i \tag{11}$$

$$\text{where } \overline{\delta}(i) \quad = \quad \beta_c e^{-2^d} + 2^{2d-3}(1 - \frac{\theta}{6}). \tag{12}$$

Since $0 = \lim_{i \mapsto \infty} \frac{\theta i}{\overline{w}(i)}$ and $1 = \lim_{i \mapsto \infty} \frac{\overline{w}(i)}{w(i)}$, it follows from (11) that

$$\limsup_{i \mapsto \infty} \overline{\rho}(i) \quad = \quad 1 - \liminf_{i \mapsto \infty} \frac{\overline{\delta}(i)}{\overline{w}(i)} = 1 - \liminf_{i \mapsto \infty} \frac{\overline{\delta}(i)}{w(i)},$$

$$\liminf_{i\mapsto\infty}\overline{\rho}(i) \quad = \quad 1-\limsup_{i\mapsto\infty}\frac{\overline{\delta}(i)}{\overline{w}(i)} = 1-\limsup_{i\mapsto\infty}\frac{\overline{\delta}(i)}{w(i)}.$$

The ratio $\frac{\overline{\delta}(i)}{w(i)}$ can be expressed in terms of $x=\frac{1}{\beta_c}$ and $z=2^d$ by:

$$\frac{\overline{\delta}(i)}{w(i)} = k(x,z) = \frac{e^{-z}}{1+z}+\frac{xz}{8}(1-\frac{\theta xz}{3}). \tag{13}$$

For $x<1$ fixed, function $k(x,z)$ is uni-modal in the interval $1\le z\le\frac{1}{x}$: it is exponentially decreasing from $k(x,1)=\frac{1}{2e}+\theta x$ to some infinitesimal $k_{min}=k(x,z_x)=O(x\log(x))$ in the interval $1\le z\le z_x=O(\log(x))$. It is (almost linearly) increasing from k_{min} to $k(x,\frac{1}{x})<\frac{1}{8}<k(x,1)$ in $z_x\le z\le\frac{1}{x}$, hence

$$0=\liminf_{i\mapsto\infty}\frac{\overline{\delta}(i)}{w(i)} \text{ and } \frac{1}{e}=\limsup_{i\mapsto\infty}\frac{\overline{\delta}(i)}{w(i)}.$$

By a similar evaluation through Lemma 1, the average BDD/BMD has size

$$\begin{aligned} a(i) &= w(i)-\delta(i)-\theta i \\ \text{where } \delta(i) &= \beta_c' e^{-2^d}+2^{2d-3}(1-\frac{\theta}{6}) = \overline{\delta}(i)-\beta_{c-1}e^{-2^d}. \end{aligned}$$

Equality $\frac{\delta(i)}{w(i)}=\frac{\overline{\delta}(i)}{w(i)}-\frac{\beta_{c-1}e^{-2^d}}{w(i)}=\frac{\overline{\delta}(i)}{w(i)}-\theta/\beta_{c-1}$ implies that the limits as $i\mapsto\infty$ are equal to the above for $\frac{\overline{\delta}(i)}{w(i)}$, and (8,9) are proved. Limit (7) follows from:

$$\frac{\overline{a}(i)}{a(i)}=\frac{\overline{w}(i)-\overline{\delta}(i)-\theta i}{w(i)-\delta(i)}=1+\frac{\beta_{c-1}(\theta+e^{-2^d})}{w(i)-\delta(i)}=1+O(\frac{1}{\beta_{c-1}}).$$

The limit (10) is established by proving that $0=\lim_{c\mapsto\infty}2^{-c}S(c)$ for the sum

$$S(c)=\sum_{1\le d\le 2^c}\frac{\overline{\delta}(c+d+2^c)}{w(c+d+2^c)}=\sum_{1\le d\le 2^c}k(x_c,2^d).$$

We evaluate $S(c)$ by (13) to find that

$$0<S(c)<\sum_d\frac{e^{-2^d}}{1+2^d}+\frac{x_c}{8}\sum_{1\le d\le 2^c}2^d = 0.154\cdots+\frac{x_c}{8}(2\beta_c-3)<1$$

and the convergence of $2^{-c}S(c)$ to 0 is exponentially fast. *Q.E.D.*

References

[1] R. E. Bryant. Graph-based algorithms for boolean function manipulation. *IEEE Trans. on Computers*, 35:8:677–691, 1986.

[2] R. E. Bryant. Symbolic boolean manipulations with ordered binary decision diagrams. *ACM Comp. Surveys*, 24:293–318, 1992.

[3] R. E. Bryant and Y.-A. Chen. Verification of arithmetic functions with binary moment diagrams. *Design Automation Conf.*, pages 535–541, 1995.

[4] D. E. Knuth. *The Art of Computer Programming, vol. 3, Sorting and Searching*. Addison Wesley, 1971.

[5] MapleSoft. *Maple 9 Guide*. Waterloo Maple Inc., 2003.

[6] I. Wegener. *The Complexity of Boolean Functions*. John Wiley and Sons Ltd, 1987.