I/O-efficient Manipulation of Binary Decision Diagrams

Steffan Christ Sølvsten

S. C. Sølvsten, J. van de Pol, A. B. Jakobsen, and M. W. B. Thomasen. *Adiar: Binary Decision Diagrams in External Memory.* 2022



Contents

What are Binary Decision Diagrams?

Why do they break?

How can we fix it?

CountPaths

Apply

Equality Checking

Contents

What are Binary Decision Diagrams?

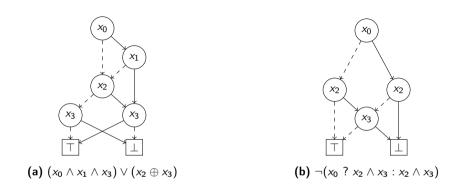
Why do they break?

How can we fix it?

CountPaths

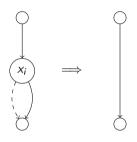
Apply

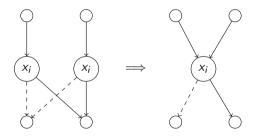
Equality Checking



Examples of (Reduced Ordered) Binary Decision Diagrams.

Theorem (Bryant '86)For a fixed variable order, if one exhaustively applies the two rules below, then one obtains the Reduced OBDD, which is a unique canonical form of the function.





(1) Remove redundant nodes

(2) Merge duplicate nodes

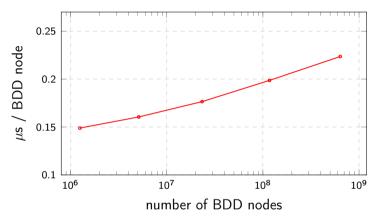
```
\begin{array}{ll} \operatorname{bdd\_apply}\left(f,\ g\ ,\ \otimes\right): \\ & \text{if}\ f,g\in\{\bot,\top\} \\ & \text{then}\ f\otimes g \\ & \text{else let}\ i = \operatorname{top}\left(f.var,\ g.var\right) \\ & t = \operatorname{bdd\_apply}\left(f[x_i:=\top],\ g[x_i:=\top],\ \otimes\right) \\ & e = \operatorname{bdd\_apply}\left(f[x_i:=\bot],\ g[x_i:=\bot],\ \otimes\right) \\ & \text{in make\_node}\left(i,t,e\right) \end{array}
```

```
\begin{aligned} \operatorname{bdd\_apply}(f,\ g,\ \otimes): \\ & \text{if}\ f,g \in \{\bot,\top\} \\ & \text{then}\ f \otimes g \\ & \text{else let}\ i = \operatorname{top}(f.var,\ g.var) \\ & t = \operatorname{bdd\_apply}(f[x_i := \top],\ g[x_i := \top],\ \otimes) \\ & e = \operatorname{bdd\_apply}(f[x_i := \bot],\ g[x_i := \bot],\ \otimes) \\ & \text{in } \operatorname{make\_node}(i,t,e) \end{aligned}
```

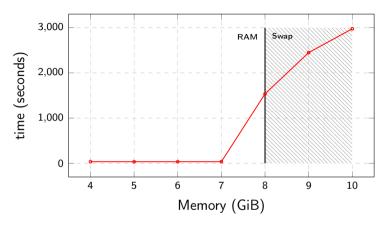
Theorem

bdd_apply runs in $O(N_f \cdot N_g)$ time.

- Memoisation (*Computation Cache*) ensures each recursion is computed only once.
- Reduction Rules can be maintained within make_node(i,t,e) in O(1) time.
 - 1 Redundancy is resolved with an if-statement.
 - 2 Duplication is avoided with a hash table (*Unique Node Table*).



Running time of *BuDDy* for the *N*-Queens problem.



Running time of BuDDy for Tic-Tac-Toe with N=21.

Contents

What are Binary Decision Diagrams?

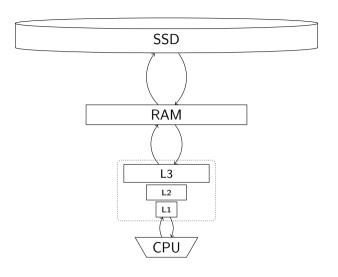
Why do they break?

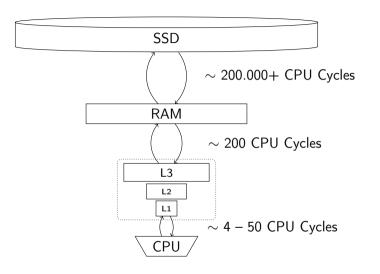
How can we fix it

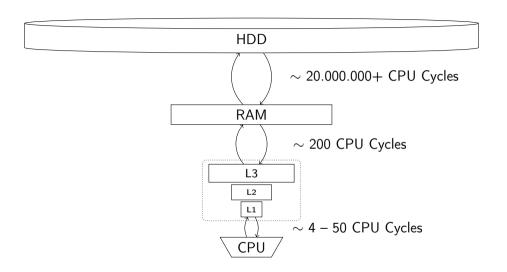
CountPaths

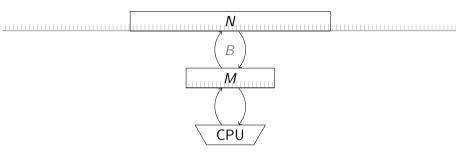
Apply

Equality Checking









The I/O model by Aggarwal and Vitter '87

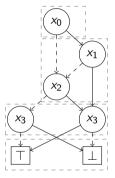
For any realistic values of N, M, and B we have that

$$N/B < \operatorname{sort}(N) \triangleq N/B \cdot \log_{M/B} N/B \ll N$$
,

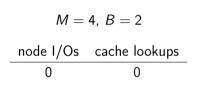
Theorem (Aggarwal and Vitter '87) N elements can be sorted in $\Theta(sort(N))$ I/Os.

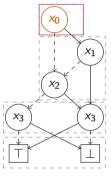
Theorem (Arge '95)

N elements can be inserted in and extracted from a Priority Queue in $\Theta(sort(N))$ I/Os.

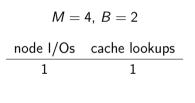


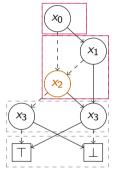
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$



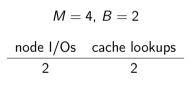


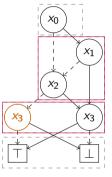
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$





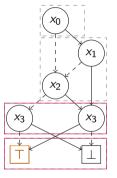
(a)
$$(x_0 \land x_1 \land x_3) \lor (x_2 \oplus x_3)$$



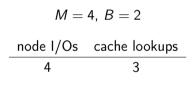


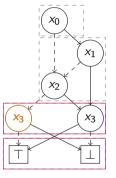
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
3 3

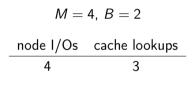


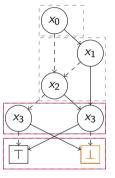
(a)
$$(x_0 \land x_1 \land x_3) \lor (x_2 \oplus x_3)$$



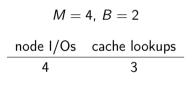


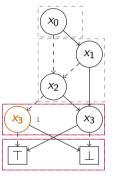
(a)
$$(x_0 \land x_1 \land x_3) \lor (x_2 \oplus x_3)$$





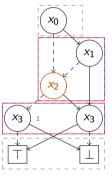
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$





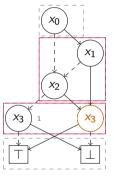
(a)
$$(x_0 \land x_1 \land x_3) \lor (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
4 3

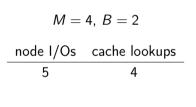


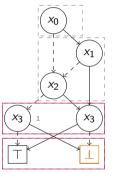
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
$$5 3$$

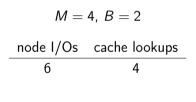


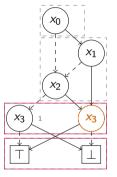
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$



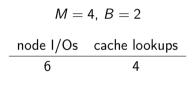


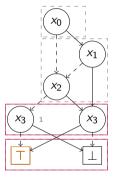
(a)
$$(x_0 \land x_1 \land x_3) \lor (x_2 \oplus x_3)$$



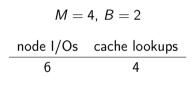


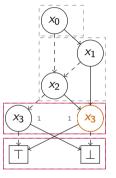
(a)
$$(x_0 \land x_1 \land x_3) \lor (x_2 \oplus x_3)$$





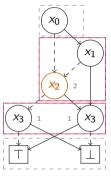
(a)
$$(x_0 \land x_1 \land x_3) \lor (x_2 \oplus x_3)$$





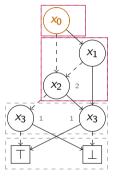
(a)
$$(x_0 \land x_1 \land x_3) \lor (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
6 4



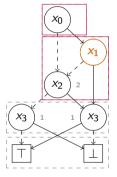
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
7 4



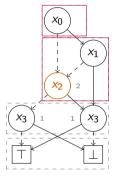
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
8 4



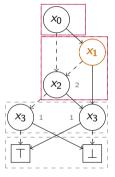
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
8 5



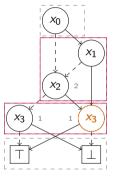
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
8 6



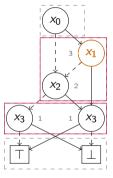
(a)
$$(x_0 \land x_1 \land x_3) \lor (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
8 6



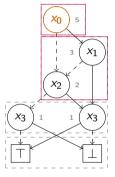
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
 9
 7

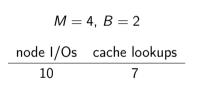


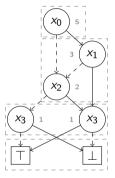
(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups
 9
 7



(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$





(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$

$$M = 4$$
, $B = 2$
node I/Os cache lookups

10 7

Algorithm	Time Complexity
bdd_pathcount	$O(N_f)$
bdd_not	$O(N_f)$
bdd_restrict	$O(N_f)$
bdd_apply	$O(N_f \cdot N_g)$
bdd_equal	O(1)

Algorithm	I/O-Complexity	
bdd_pathcount	$O(N_f)$	
bdd_not	$O(N_f)$	
bdd_restrict	$O(N_f)$	
bdd_apply	$O(N_f \cdot N_g)$	
bdd_equal	O(1)	

Contents

What are Binary Decision Diagrams?

Why do they break?

How can we fix it?

CountPaths

Apply

Equality Checking

Contents

What are Binary Decision Diagrams?

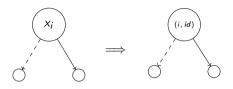
Why do they break?

How can we fix it?

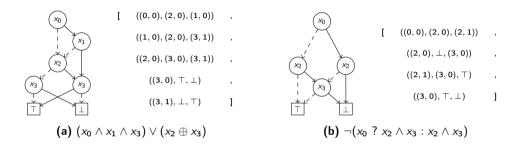
CountPaths

Apply

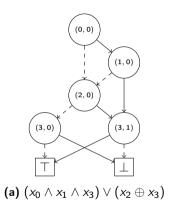
Equality Checking

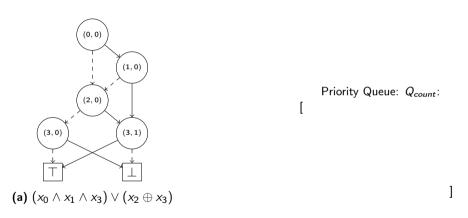


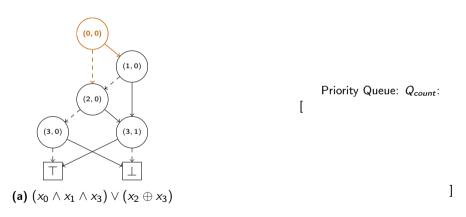
$$(i_1, id_1) < (i_2, id_2) \equiv i_1 < i_2 \lor (i_1 = i_2 \land id_i < id_j)$$

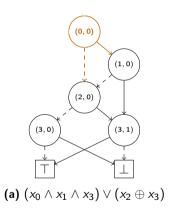


Node-based representation of prior shown $\ensuremath{\mathsf{BDDs}}$

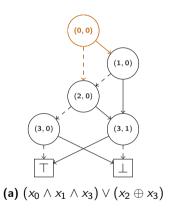


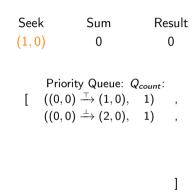


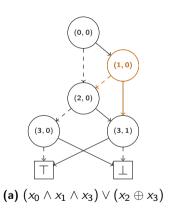


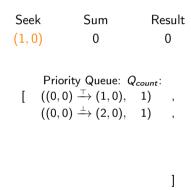


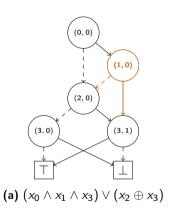
Priority Queue:
$$Q_{count}$$
: [$((0,0) \xrightarrow{\top} (1,0), 1)$, $((0,0) \xrightarrow{\bot} (2,0), 1)$,

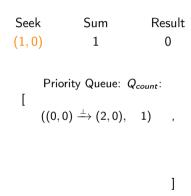


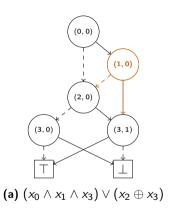


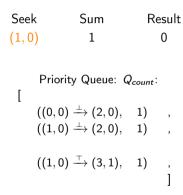


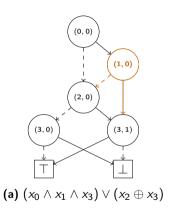


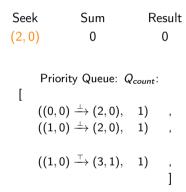


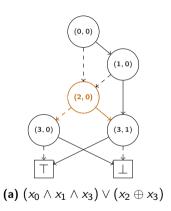


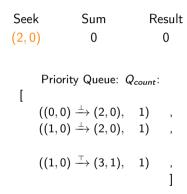


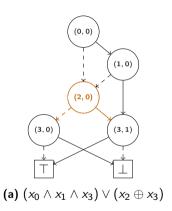


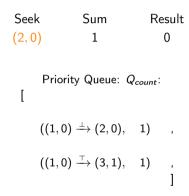


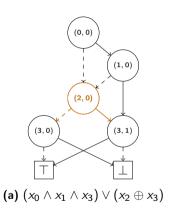


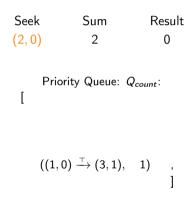


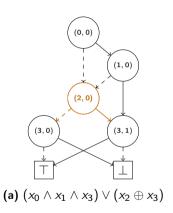




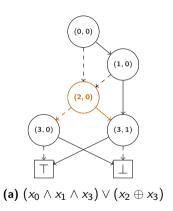




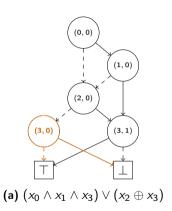




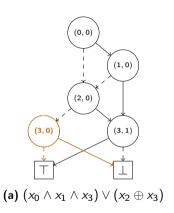
```
Seek
                                   Sum
                                                                      Result
(2,0)
                                        2
                                                                             0
              Priority Queue: Qcount:
             \begin{array}{cccc} ((2,0) \xrightarrow{\bot} (3,0), & 2) & , \\ ((1,0) \xrightarrow{\top} (3,1), & 1) & , \\ ((2,0) \xrightarrow{\top} (3,1), & 2) & ] \end{array}
```

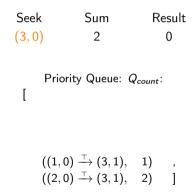


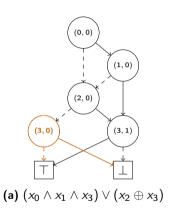
```
Seek
                                    Sum
                                                                      Result
(3,0)
                                        0
                                                                             0
              Priority Queue: Qcount:
             \begin{array}{cccc} ((2,0) \xrightarrow{\bot} (3,0), & 2) & , \\ ((1,0) \xrightarrow{\top} (3,1), & 1) & , \\ ((2,0) \xrightarrow{\top} (3,1), & 2) & ] \end{array}
```

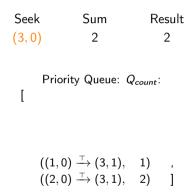


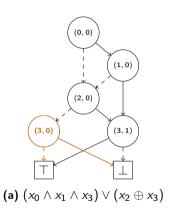
```
Seek
                                    Sum
                                                                      Result
(3,0)
                                        0
                                                                             0
              Priority Queue: Qcount:
             \begin{array}{cccc} ((2,0) \xrightarrow{\bot} (3,0), & 2) & , \\ ((1,0) \xrightarrow{\top} (3,1), & 1) & , \\ ((2,0) \xrightarrow{\top} (3,1), & 2) & ] \end{array}
```

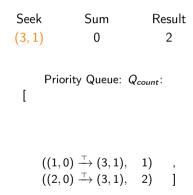


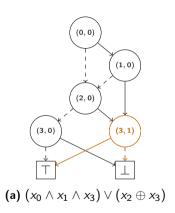


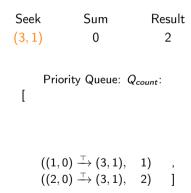


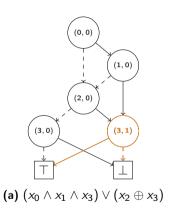


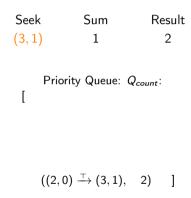


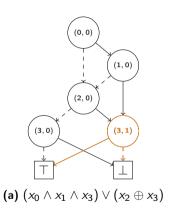


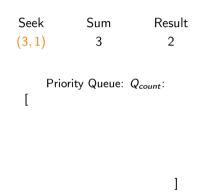


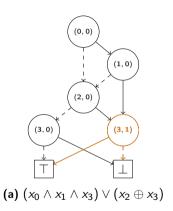


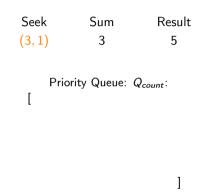


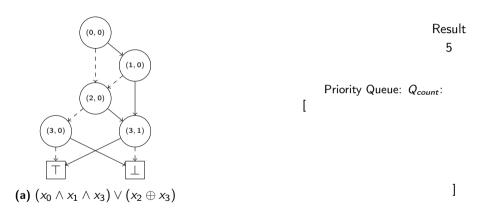












Contents

What are Binary Decision Diagrams?

Why do they break?

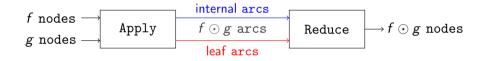
How can we fix it?

CountPaths

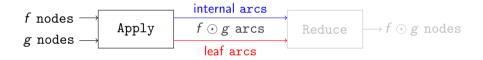
Apply

Equality Checking

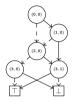
Apply



Apply



Apply: Example

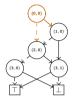


(a) $(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$



(b)
$$\neg (x_0 ? x_2 \lor x_3 : x_2 \land x_3)$$

(c) $(a) \wedge (b)$

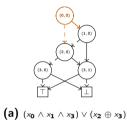


(a) $(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$



(b) $\neg (x_0 ? x_2 \lor x_3 : x_2 \land x_3)$

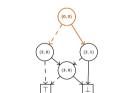
(c) $(a) \wedge (b)$



Priority Queue: Q_{app:1}:

- [$(0,0) \xrightarrow{\top} ((1,0),(2,1))$,
 - $(0,0) \xrightarrow{\perp} ((2,0),(2,0))$

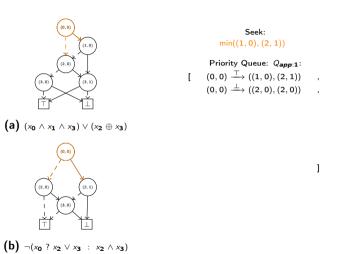
(0,0)



]

(c) $(a) \wedge (b)$

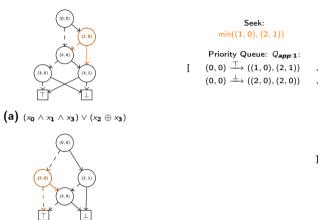
(b) $\neg (x_0 ? x_2 \lor x_3 : x_2 \land x_3)$



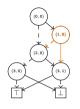
(0,0)

(c) $(a) \wedge (b)$

(b) $\neg (x_0 ? x_2 \lor x_3 : x_2 \land x_3)$



(c) (a) ∧ (b)



(a)
$$(x_0 \wedge x_1 \wedge x_3) \vee (x_2 \oplus x_3)$$



(b) $\neg (x_0 ? x_2 \lor x_3 : x_2 \land x_3)$

Seek: min((1,0),(2,1))

Priority Queue: Qapp:1:

 $(0,0) \xrightarrow{\top} ((1,0),(2,1))$

 $(0,0) \xrightarrow{\perp} ((2,0),(2,0))$

 $(1,0) \xrightarrow{\perp} ((2,0),(2,1))$

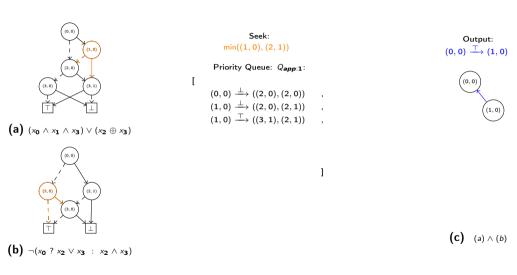
 $(1,0) \xrightarrow{\top} ((3,1),(2,1))$

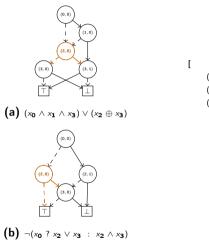
(0,0)

(1,0)

1

(c) $(a) \wedge (b)$



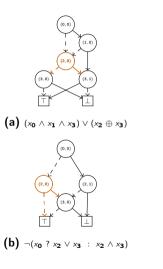


Seek: min((2,0),(2,0))Priority Queue: Qapp:1: $(0,0) \xrightarrow{\perp} ((2,0),(2,0))$ $(1,0) \xrightarrow{\perp} ((2,0),(2,1))$ $(1,0) \xrightarrow{\top} ((3,1),(2,1))$

Output:



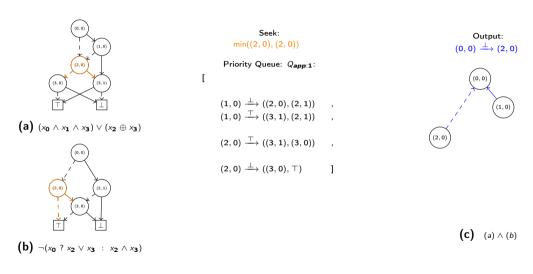
(c) (a) ∧ (b)

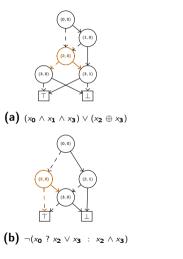


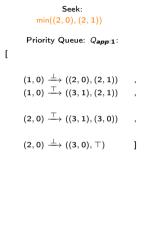
Seek: min((2,0),(2,0))Priority Queue: Qapp:1: $(0,0) \xrightarrow{\perp} ((2,0),(2,0))$ $(1,0) \xrightarrow{\perp} ((2,0),(2,1))$ $(1,0) \xrightarrow{\top} ((3,1),(2,1))$ $(2,0) \xrightarrow{\top} ((3,1),(3,0))$ $(2,0) \xrightarrow{\perp} ((3,0),\top)$]

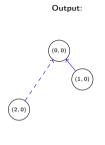
Output: (2,0)

(c) $(a) \wedge (b)$

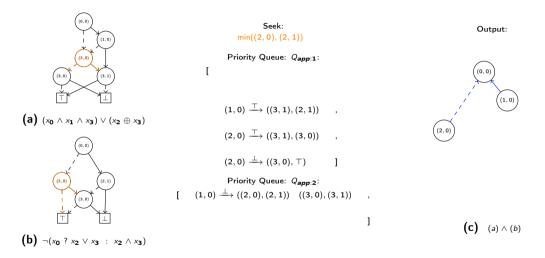


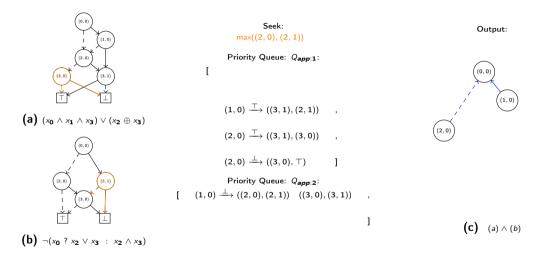


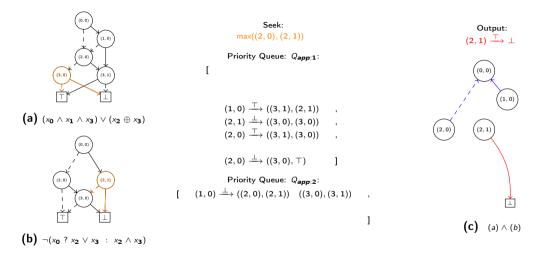


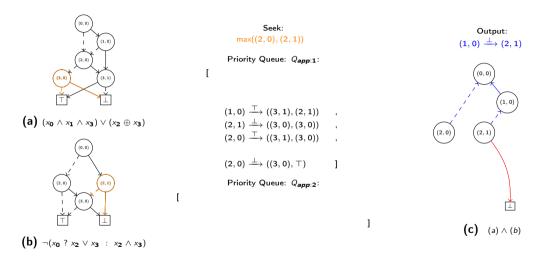


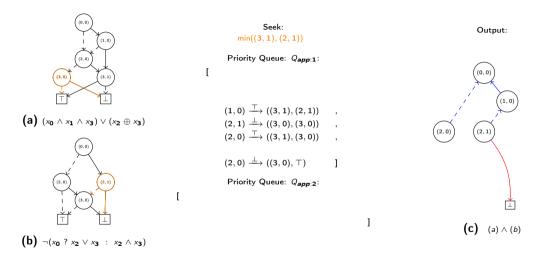
(c) $(a) \wedge (b)$

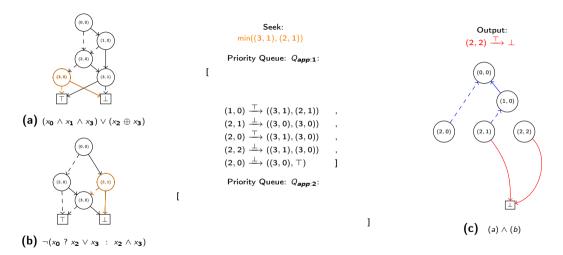


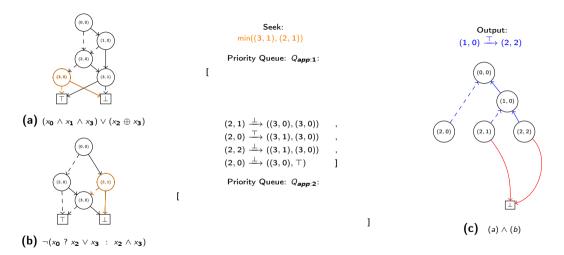


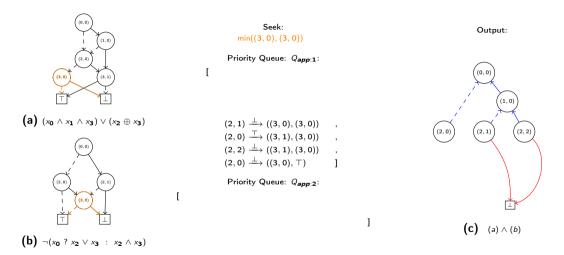


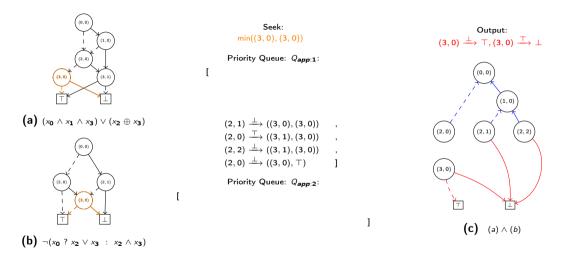


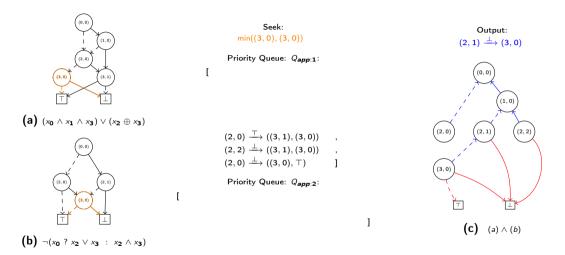


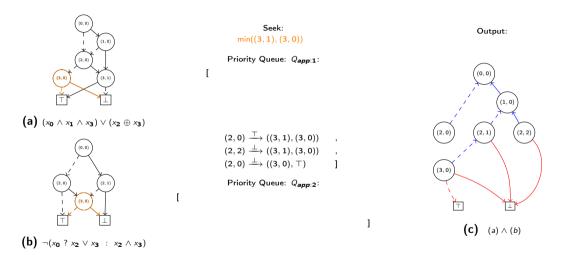


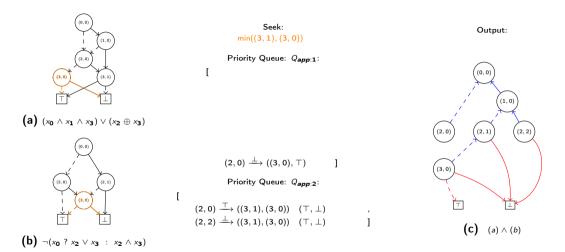


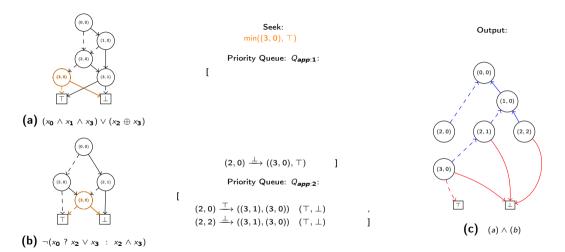


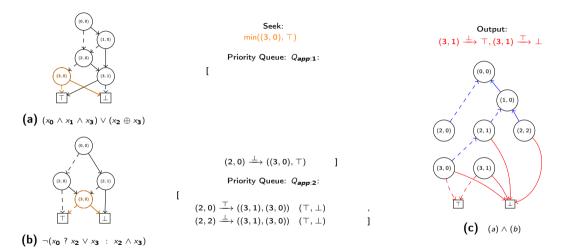


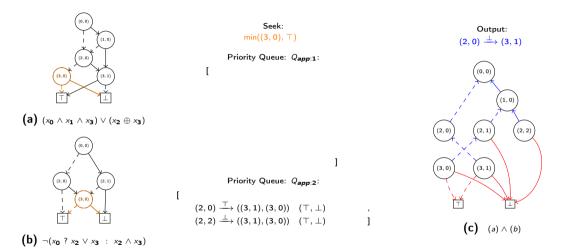


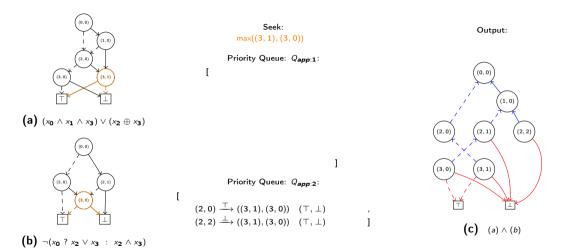


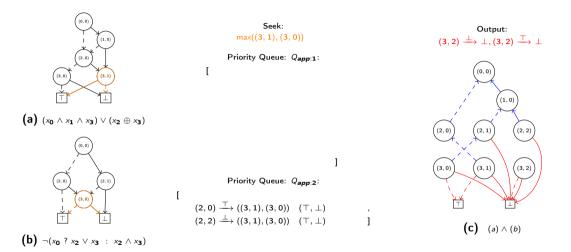


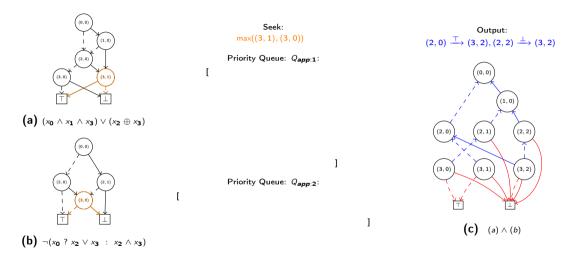


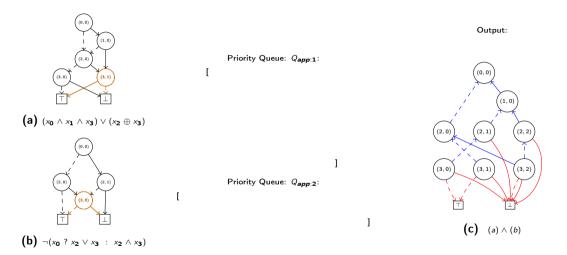




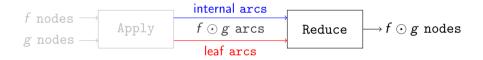


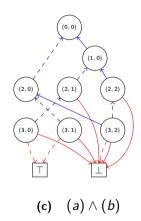


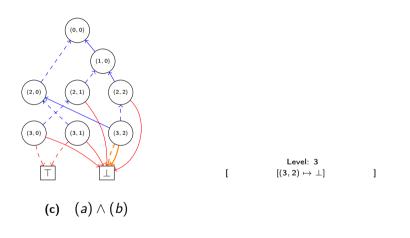


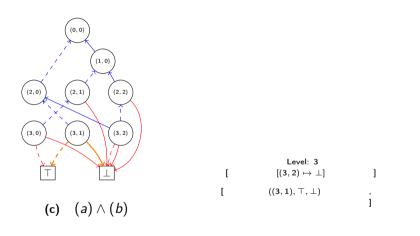


Apply

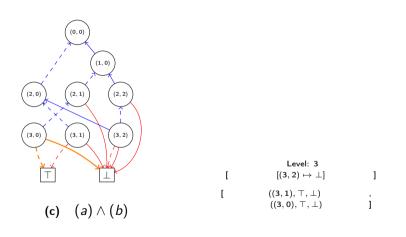


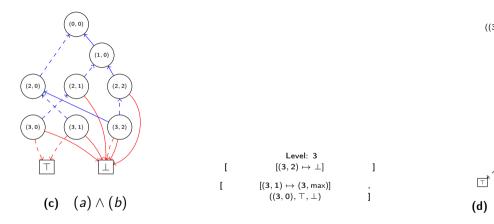




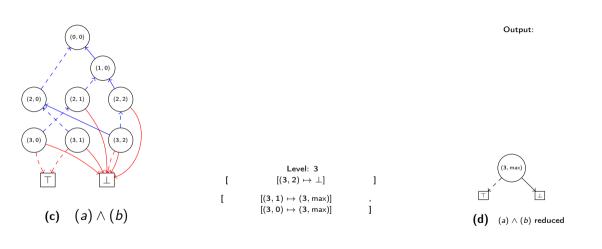


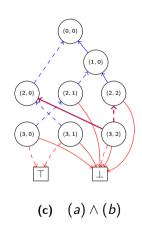
17

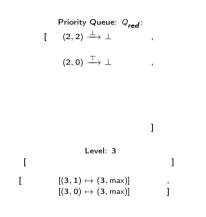






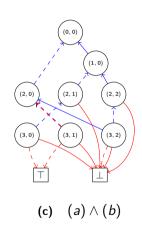


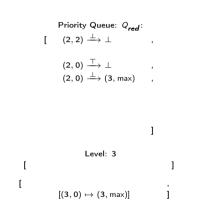






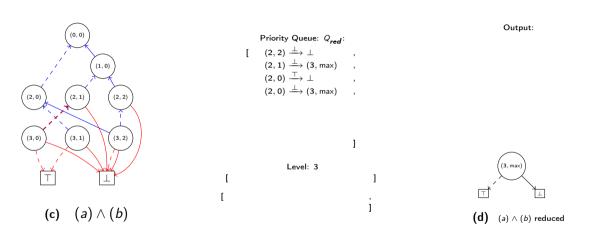


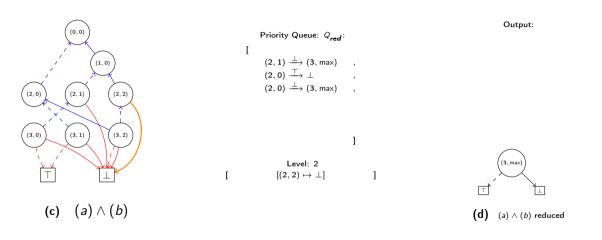


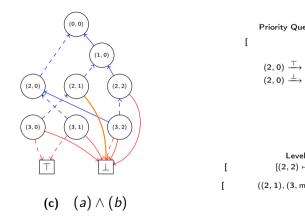


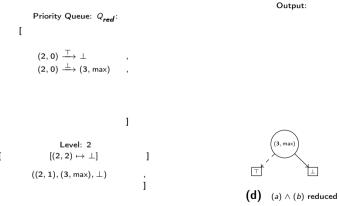


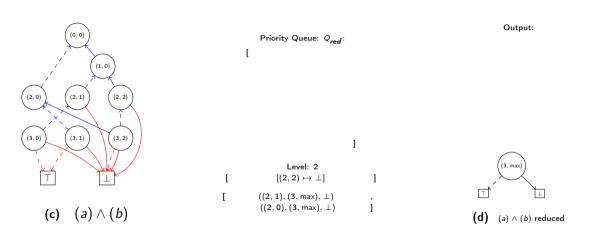


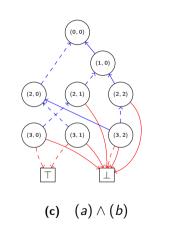


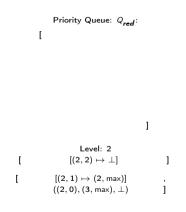


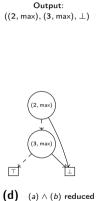


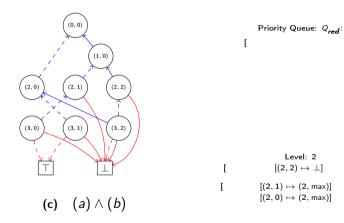


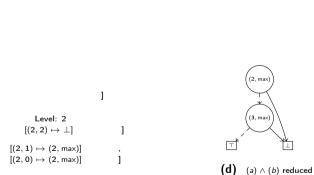


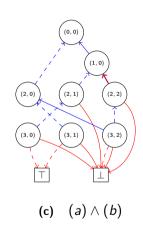


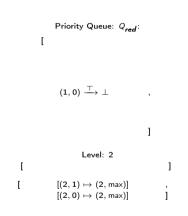


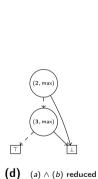


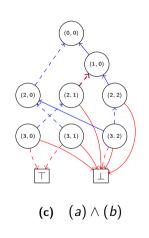


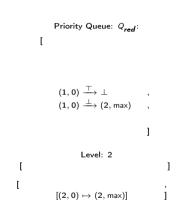


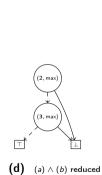


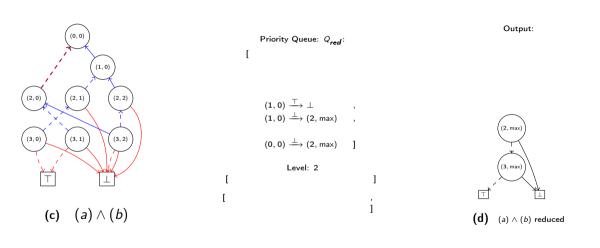


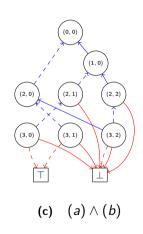


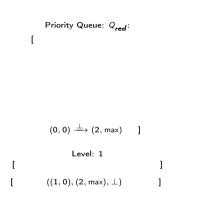


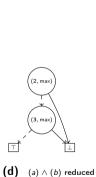


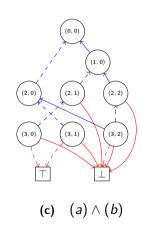


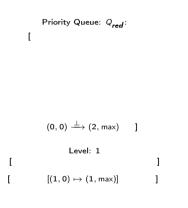


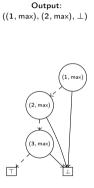


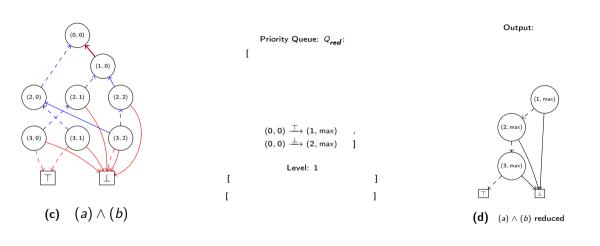


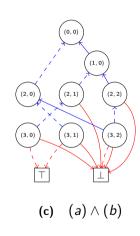


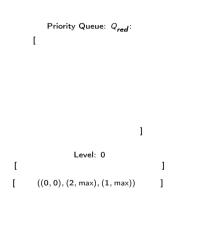


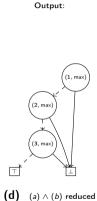


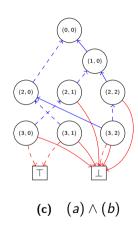


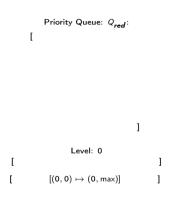


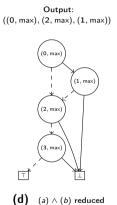


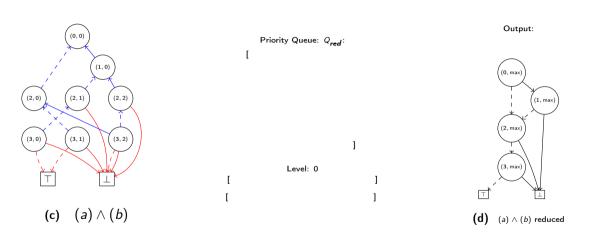




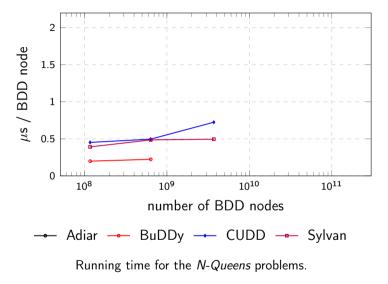


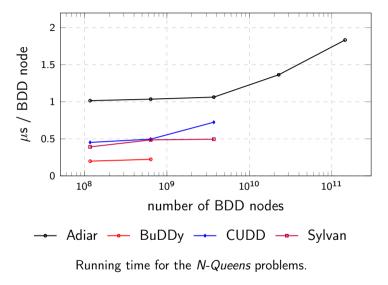






Algorithm	I/O-Complexity
bdd_pathcount	$O(\operatorname{sort}(N_f))$
bdd_not	$O(N_f/B)$
bdd_restrict	$O(sort(N_f))$
bdd_apply	$O(\operatorname{sort}(N_f \cdot N_g))$





Contents

What are Binary Decision Diagrams?

Why do they break?

How can we fix it?

CountPaths

Apply

Equality Checking

Algorithm	I/O-Complexity	
bdd_pathcount	$O(\operatorname{sort}(N_f))$	
bdd_not	$O(N_f/B)$	
bdd_restrict	$O(\operatorname{sort}(N_f))$	
bdd_apply	$O(\operatorname{sort}(N_f \cdot N_g))$	

Algorithm	I/O-Complexity	
bdd_pathcount	$O(\operatorname{sort}(N_f))$	
bdd_not	$O(N_f/B)$	
bdd_restrict	$O(\operatorname{sort}(N_f))$	
bdd_apply	$O(\operatorname{sort}(N_f \cdot N_g))$	
bdd_equal	?	

$$f\leftrightarrow g\equiv \top$$

$$f \leftrightarrow g \equiv \top$$

$$\underbrace{O(\mathsf{sort}(\mathit{N}^2))}_{\mathsf{Apply}} + \underbrace{O(\mathsf{sort}(\mathit{N}^2))}_{\mathsf{Reduce}} + \underbrace{O(1))}_{\mathsf{check is }\top} = O(\mathsf{sort}(\mathit{N}^2))$$

Theorem (Bryant '86)

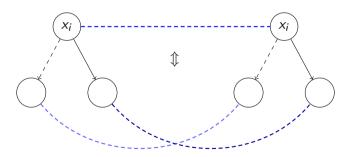
Theorem (Bryant '86)

Let π be a variable order and $f: \mathbb{B}^n \to \mathbb{B}$ then there exists a unique (up to isomorphism) Reduced Ordered Binary Decision Diagram representing f with ordering π .

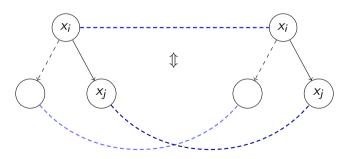
Trivial cases: $f \not\equiv g$ if there is a mismatch in

	$N_f eq N_g$	Number of nodes	O(1) I/Os
•	$L_f eq L_g$	Number of levels	<i>O</i> (1) I/Os
•	$N_{f,i} \neq N_{g,i}$	Number of nodes on a level	O(L/B) I/Os
•	$L_{f,i} \neq L_{g,i}$	Label of an <i>i</i> th level	O(L/B) I/Os

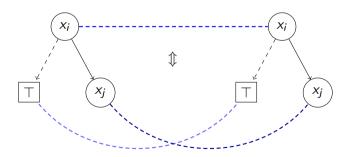
Theorem (Bryant '86)



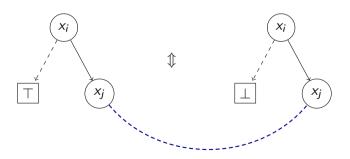
Theorem (Bryant '86)



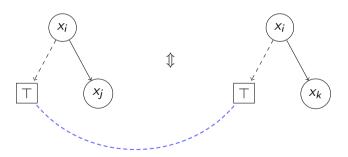
Theorem (Bryant '86)



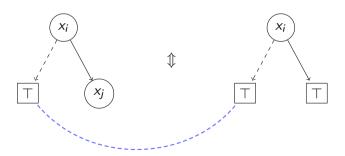
Theorem (Bryant '86)



Theorem (Bryant '86)



Theorem (Bryant '86)



Theorem (Bryant '86)

Let π be a variable order and $f: \mathbb{B}^n \to \mathbb{B}$ then there exists a unique (up to isomorphism) Reduced Ordered Binary Decision Diagram representing f with ordering π .

IsIsomorphic(f, g)

- Check whether root v_f of f and root v_g of g have a local violation.
- Check $low(v_f) \sim low(v_g)$ and $high(v_f) \sim high(v_g)$ "recursively".

Return false on first violation. If there are no violations then return true.

Theorem (Bryant '86)

Let π be a variable order and $f: \mathbb{B}^n \to \mathbb{B}$ then there exists a unique (up to isomorphism) Reduced Ordered Binary Decision Diagram representing f with ordering π .

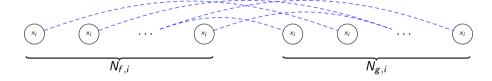
IsIsomorphic(f, g)

- Check whether root v_f of f and root v_g of g have a local violation.
- Check $low(v_f) \sim low(v_g)$ and $high(v_f) \sim high(v_g)$ "recursively".

Return false on first violation. If there are no violations then return true.

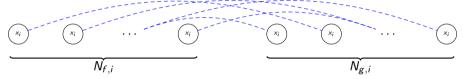
$$\underbrace{O(\mathsf{sort}(\mathit{N}^2))}_{\mathtt{Apply'}} + \underbrace{O(\mathsf{sort}(\mathit{N}^2))}_{\mathtt{Reduce}} + \underbrace{O(1))}_{\mathtt{check is }\top} = O(\mathsf{sort}(\mathit{N}^2))$$

Theorem (Bryant '86)



Theorem (Bryant '86)

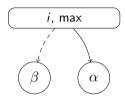
Let π be a variable order and $f: \mathbb{B}^n \to \mathbb{B}$ then there exists a unique (up to isomorphism) Reduced Ordered Binary Decision Diagram representing f with ordering π .



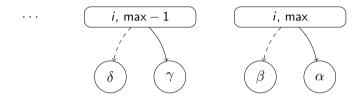
Return false if more than $N_{f,i} = N_{g,i}$ pairs of nodes are checked on level i.

$$\underbrace{\mathit{O}(\mathsf{sort}(\Sigma_i \; \mathit{N}_{f,i}))}_{\mathsf{Apply''}} = \mathit{O}(\mathsf{sort}(\mathit{N}))$$

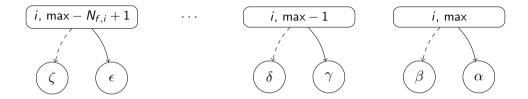
Observation



Observation



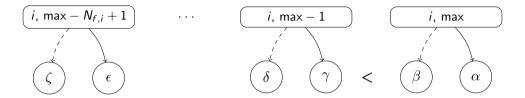
Observation



Observation

Each level output by the Reduce algorithm has the following properties:

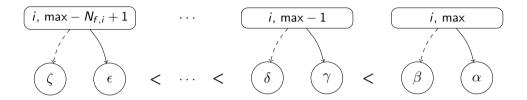
■ Nodes on level *i* have their identifiers *consecutively* numbered.



Observation

Each level output by the Reduce algorithm has the following properties:

■ Nodes on level *i* have their identifiers *consecutively* numbered.



Observation

- Nodes on level *i* have their identifiers *consecutively* numbered.
- Nodes on level *i* are output sorted by their children.

Theorem

If G_f and G_g are outputs of Reduce.

 $G_f \sim G_g \iff \textit{For all } i \in [0; N_f) \textit{ the node } G_f[i] \textit{ matches } G_g[i] \textit{ numerically.}$

Proof.

← : Must describe the exact same graph.

 \Rightarrow : Strong induction on BDD levels bottom-up.

Theorem

If G_f and G_g are outputs of Reduce.

 $G_f \sim G_g \iff For \ all \ i \in [0; N_f) \ the \ node \ G_f[i] \ matches \ G_g[i] \ numerically.$

Proof.

← : Must describe the exact same graph.

 \Rightarrow : Strong induction on BDD levels bottom-up.

Corollary

If G_f and G_g are outputs of Reduce then $f \equiv g$ is computable using $2 \cdot N/B$ I/Os.

$$\begin{array}{c|c} & \text{Algorithm} & \text{Time (s)} \\ \hline f \leftrightarrow g \equiv \top & 0.38 \end{array}$$

Checking the (EPFL Benchmark) voter circuit's single output gate ($|N_f| = |N_g| = 5.76$ MiB).

Algorithm Time (s)
$$f \leftrightarrow g \equiv \top$$
 0.38
 $O(\operatorname{sort}(N))$ 0.058

Checking the (EPFL Benchmark) *voter* circuit's single output gate ($|N_f| = |N_g| = 5.76$ MiB).

Algorithm	Time (s)	
$f\leftrightarrow g\equiv \top$	0.38	
$O(\operatorname{sort}(N))$	0.058	
2N/B	0.006	

Checking the (EPFL Benchmark) voter circuit's single output gate ($|N_f| = |N_g| = 5.76$ MiB).

Contents

What are Binary Decision Diagrams?
Why do they break?
How can we fix it?
CountPaths
Apply
Equality Checking $O(N_f) \quad O(\text{sort}(N_f)) \quad O(\text{sort}(N_f \cdot N_g)) \quad O(\text{sort}(N_f \cdot N_g)) \quad O(\text{sort}(N_f \cdot N_g)) \quad O(N_f \cdot N_g) \quad O(N_f \cdot N_$