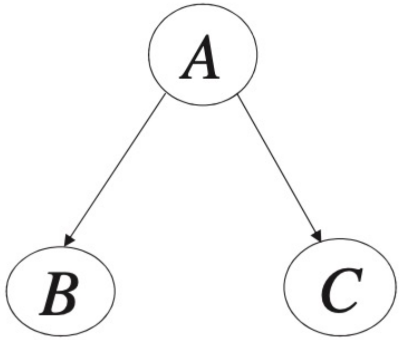
Reducing Probabilistic Reasoning (MAR) to Weighted Model Counting (WMC)

Probabilistic Reasoning



A	Θ_A
a_1	[illegible]
a_2	[illegible]

A	B	$\Theta_{B\|A}$
a_1	b_1	.1
a_1	b_2	.9
a_2	b_1	.2
a_2	b_2	.8

A	C	$\Theta_{C\|A}$
a_1	c_1	.1
a_1	c_2	.9
a_2	c_1	.2
a_2	c_2	.8

Network instantiation	
$a_1b_1c_1$	$.1 \cdot .1 \cdot .1 = .001$
$a_1b_1c_2$	$.1 \cdot .1 \cdot .9 = .009$
$a_1b_2c_1$	$.1 \cdot .9 \cdot .1 = .009$
$a_1b_2c_2$	$.1 \cdot .9 \cdot .9 = .081$
$a_2b_1c_1$	$.9 \cdot .2 \cdot .2 = .036$
$a_2b_1c_2$	$.9 \cdot .2 \cdot .8 = .144$
$a_2b_2c_1$	$.9 \cdot .8 \cdot .2 = .144$
$a_2b_2c_2$	$.9 \cdot .8 \cdot .8 = .576$

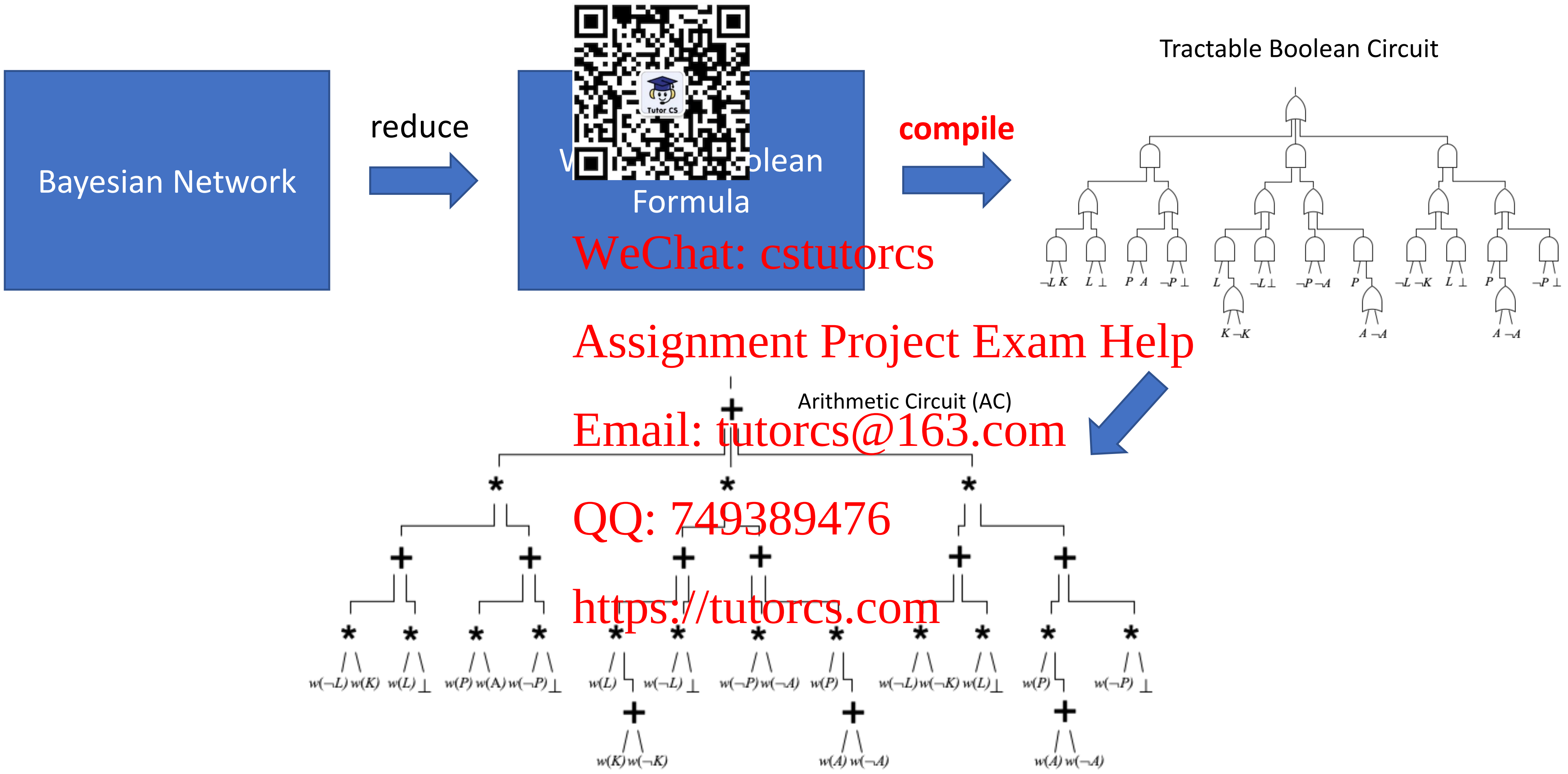
MPE: Find most likely instantiation that satisfies a given event (e.g., B=b_2)

MAR: Sum probabilities of instantiations that satisfy a given event (e.g., B=b_2)

Distribution

Factor

The Pipeline



Weighted Model Counting (WMC)

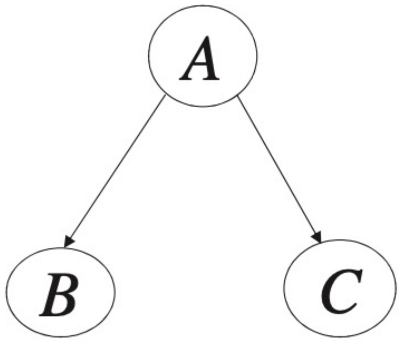
Definition 11.3 (Weighted model counting, WMC). Let Δ be a propositional sentence over Boolean variables $X_1, \ldots, X_n$ and let Wt be a function that assigns a weight $Wt(x_i) \geq 0$ to each value x_i of variable X_i. The *weighted model count* of Δ is defined as the sum of weights assigned to its models:

$$\mathrm{WMC}(\Delta) \stackrel{def}{=} \sum_{x_1,\ldots,x_n \models \Delta} Wt(x_1, \ldots, x_n),$$

where

$$Wt(x_1, \ldots, x_n) \stackrel{def}{=} \prod_{i=1}^{n} Wt(x_i).$$

Reduction: MAR → WMC



A	Θ_A
a_1	.1
a_2	.9

A	B	$\Theta_{B\|A}$
a_1	b_1	.1
a_1	b_2	.9
a_2	b_1	.2
a_2	b_2	.8

A	C	$\Theta_{C\|A}$
a_1	c_1	.1
a_1	c_2	.9
a_2	c_1	.2
a_2	c_2	.8

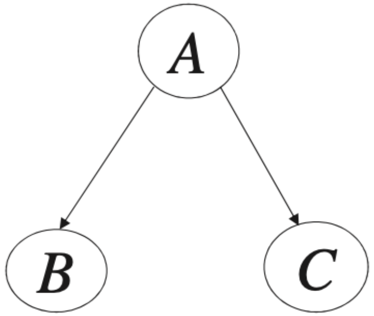
Indicator Variables (Boolean)

$$I_{a_1}, I_{a_2}, \quad I_{b_1}, I_{b_2}, \quad I_{c_1}, I_{c_2},$$

Parameter Variables (Boolean)

$$P_{a_1}, P_{a_2}, \quad P_{b_1|a_1}, P_{b_2|a_1}, P_{b_1|a_2}, P_{b_2|a_2}, \quad P_{c_1|a_1}, P_{c_2|a_1}, P_{c_1|a_2}, P_{c_2|a_2}.$$

Reduction: MAR → WMC



A	Θ_A
a_1	.1
a_2	.9

A	B	$\Theta_{B\|A}$
a_1	b_1	.1
a_1	b_2	.9
a_2	b_1	.2
a_2	b_2	.8

A	C	$\Theta_{C\|A}$
a_1	c_1	.1
a_1	c_2	.9
a_2	c_1	.2
a_2	c_2	.8

variable X with values $x_1, x_2, \ldots, x_k$

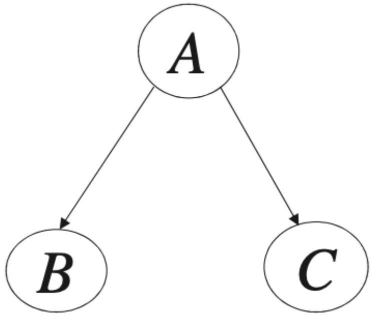
$$I_{x_1} \vee I_{x_2} \vee \ldots \vee I_{x_k}$$

$$\neg I_{x_i} \vee \neg I_{x_j}, \quad \text{for } i < j.$$

$$I_{u_1} \wedge I_{u_2} \wedge \ldots \wedge I_{u_m} \wedge I_x \iff P_{x|u_1,u_2,\ldots,u_m}.$$

$$Wt(I_x) = Wt(\neg I_x) = Wt(\neg P_{x|\mathbf{u}}) = 1 \;\; and \;\; Wt(P_{x|\mathbf{u}}) = \theta_{x|\mathbf{u}}.$$

Reduction: MAR → WMC



A	Θ_A
a_1	.1
a_2	.9

A	B	$\Theta_{B\|A}$
a_1	b_1	.1
a_1	b_2	.9
a_2	b_1	.2
a_2	b_2	.8

A	C	$\Theta_{C\|A}$
a_1	c_1	.1
a_1	c_2	.9
a_2	c_1	.2
a_2	c_2	.8

variable X with values $x_1, x_2, \ldots, x_k$

$$I_{x_1} \vee I_{x_2} \vee \ldots \vee I_{x_k}$$

$$\neg I_{x_i} \vee \neg I_{x_j}, \qquad \text{for } i < j.$$

$$I_{u_1} \wedge I_{u_2} \wedge \ldots \wedge I_{u_m} \wedge I_x \iff P_{x|u_1,u_2,\ldots,u_m}.$$

	or Clauses	
	$I_{a_1} \vee I_{a_2}$	$\neg I_{a_1} \vee \neg I_{a_2}$
B	$I_{b_1} \vee I_{b_2}$	$\neg I_{b_1} \vee \neg I_{b_2}$
C	$I_{c_1} \vee I_{c_2}$	$\neg I_{c_1} \vee \neg I_{c_2}$

	Parameter Clauses	
A	$I_{a_1} \iff P_{a_1}$	$I_{a_2} \iff P_{a_2}$
B	$I_{a_1} \wedge I_{b_1} \iff P_{b_1\|a_1}$	$I_{a_1} \wedge I_{b_2} \iff P_{b_2\|a_1}$
	$I_{a_2} \wedge I_{b_1} \iff P_{b_1\|a_2}$	$I_{a_2} \wedge I_{b_2} \iff P_{b_2\|a_2}$
C	$I_{a_1} \wedge I_{c_1} \iff P_{c_1\|a_1}$	$I_{a_1} \wedge I_{c_2} \iff P_{c_2\|a_1}$
	$I_{a_2} \wedge I_{c_1} \iff P_{c_1\|a_2}$	$I_{a_2} \wedge I_{c_2} \iff P_{c_2\|a_2}$

$$Wt(I_x) = Wt(\neg I_x) = Wt(\neg P_{x|\mathbf{u}}) = 1 \ \ and \ \ Wt(P_{x|\mathbf{u}}) = \theta_{x|\mathbf{u}}.$$

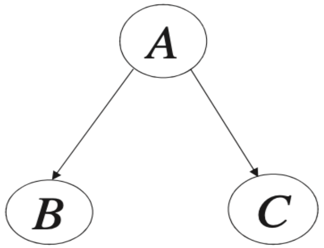
Reduction: MAR → WMC

Let $\mathcal{N}$ *be a Bayesian network inducing probability distribution* Pr
$\Delta_{\mathcal{N}}$ *be its CNF encoding*

For any evidence $\mathbf{e} = e_1, \ldots, e_k$,

$$\Pr(\mathbf{e}) = \mathrm{WMC}(\Delta_{\mathcal{N}} \wedge I_{e_1} \wedge \ldots \wedge I_{e_k}),$$

Reduction: MAR → WMC (why it works)



A	Θ_A
a_1	.1
a_2	.9

A	B	$\Theta_{B\mid A}$
a_1	b_1	.1
a_1	b_2	.9
a_2	b_1	.2
a_2	b_2	.8

A	C	$\Theta_{C\mid A}$
a_1	c_1	.1
a_1	c_2	.9
a_2	c_1	.2
a_2	c_2	.8

Indicator Clauses		
A	$I_{a_1} \vee I_{a_2}$	$\neg I_{a_1} \vee \neg I_{a_2}$
B	$I_{b_1} \vee I_{b_2}$	$\neg I_{b_1} \vee \neg I_{b_2}$
C	$I_{c_1} \vee I_{c_2}$	$\neg I_{c_1} \vee \neg I_{c_2}$

$\Delta_{\mathcal{N}}$:

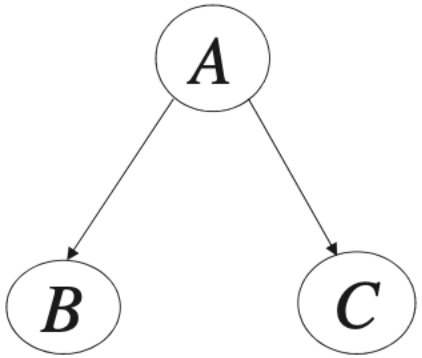
Parameter Clauses		
A	$I_{a_1} \iff P_{a_1}$	$I_{a_2} \iff P_{a_2}$
B	$I_{a_1} \wedge I_{b_1} \iff P_{b_1\mid a_1}$	$I_{a_1} \wedge I_{b_2} \iff P_{b_2\mid a_1}$
	$I_{a_2} \wedge I_{b_1} \iff P_{b_1\mid a_2}$	$I_{a_2} \wedge I_{b_2} \iff P_{b_2\mid a_2}$
C	$I_{a_1} \wedge I_{c_1} \iff P_{c_1\mid a_1}$	$I_{a_1} \wedge I_{c_2} \iff P_{c_2\mid a_1}$
	$I_{a_2} \wedge I_{c_1} \iff P_{c_1\mid a_2}$	$I_{a_2} \wedge I_{c_2} \iff P_{c_2\mid a_2}$

	Truth assignment sets these variables to true and all others to false	Weight of truth assignment
$a_1b_1c_1$	$\omega_0 : I_{a_1}\ I_{b_1}\ I_{c_1}\ P_{a_1}\ P_{b_1\mid a_1}\ P_{c_1\mid a_1}$	$.1 \cdot .1 \cdot .1 = .001$
$a_1b_1c_2$	$\omega_1 : I_{a_1}\ I_{b_1}\ I_{c_2}\ P_{a_1}\ P_{b_1\mid a_1}\ P_{c_2\mid a_1}$	$.1 \cdot .1 \cdot .9 = .009$
$a_1b_2c_1$	$\omega_2 : I_{a_1}\ I_{b_2}\ I_{c_1}\ P_{a_1}\ P_{b_2\mid a_1}\ P_{c_1\mid a_1}$	$.1 \cdot .9 \cdot .1 = .009$
$a_1b_2c_2$	$\omega_3 : I_{a_1}\ I_{b_2}\ I_{c_2}\ P_{a_1}\ P_{b_2\mid a_1}\ P_{c_2\mid a_1}$	$.1 \cdot .9 \cdot .9 = .081$
$a_2b_1c_1$	$\omega_4 : I_{a_2}\ I_{b_1}\ I_{c_1}\ P_{a_2}\ P_{b_1\mid a_1}\ P_{c_1\mid a_2}$	$.9 \cdot .2 \cdot .2 = .036$
$a_2b_1c_2$	$\omega_5 : I_{a_2}\ I_{b_1}\ I_{c_2}\ P_{a_2}\ P_{b_1\mid a_1}\ P_{c_2\mid a_2}$	$.9 \cdot .2 \cdot .8 = .144$
$a_2b_2c_1$	$\omega_6 : I_{a_2}\ I_{b_2}\ I_{c_1}\ P_{a_2}\ P_{b_2\mid a_1}\ P_{c_1\mid a_2}$	$.9 \cdot .8 \cdot .2 = .144$
$a_2b_2c_2$	$\omega_7 : I_{a_2}\ I_{b_2}\ I_{c_2}\ P_{a_2}\ P_{b_2\mid a_1}\ P_{c_2\mid a_2}$	$.9 \cdot .8 \cdot .8 = .576$

$$\mathbf{e} = a_1c_2$$

$$\text{WMC}(\Delta_{\mathcal{N}} \wedge I_{a_1} \wedge I_{c_2}) = Wt(\omega_1) + Wt(\omega_3) = .009 + .081 = .09 = \Pr(\mathbf{e}).$$

Local Structure



A	θ_A
true	.5
false	.5

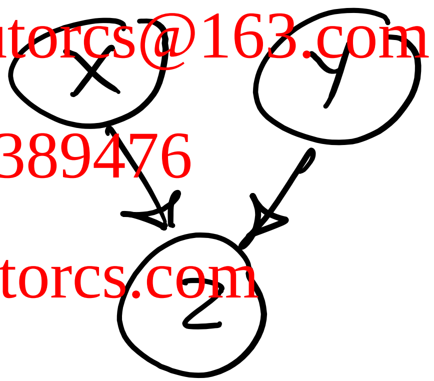
A	B	$\theta_{B\|A}$
true	true	1
true	false	0
false	true	0
false	false	1

A	C	$\theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8

--Zero/One parameters (logical constraints)

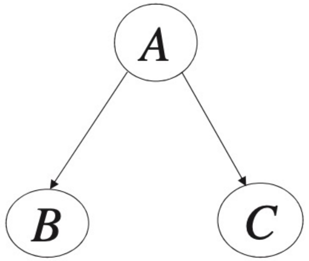
--Equal parameters

Context-Specific Independence (CSI)



X	Y	Z	Pr(Z|X,Y)
t	t	t	.8
t	t	f	.2
t	f	t	.8
t	f	f	.2
⋮			

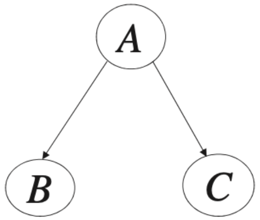
Encoding Local Structure



A	Θ_A
a_1	.1
a_2	.9

A	B	$\Theta_{B\|A}$
a_1	b_1	.1
a_1	b_2	.9
a_2	b_1	.2
a_2	b_2	.8

A	C	$\Theta_{C\|A}$
a_1	c_1	.1
a_1	c_2	.9
a_2	c_1	.2
a_2	c_2	.8



A	θ_A
true	.5
false	.5

A	B	$\theta_{B\|A}$
true	true	1
true	false	0
false	true	0
false	false	1

A	C	$\theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8

Indicator Clauses

A	$I_{a_1} \vee I_{a_2}$	$\neg I_{a_1} \vee \neg I_{a_2}$
B	$I_{b_1} \vee I_{b_2}$	$\neg I_{b_1} \vee \neg I_{b_2}$
C	$I_{c_1} \vee I_{c_2}$	$\neg I_{c_1} \vee \neg I_{c_2}$

Parameter Clauses

A	$I_{a_1} \Longleftrightarrow P_{a_1}$	$I_{a_2} \Leftrightarrow P_{a_2}$
B	$I_{a_1} \wedge I_{b_1} \Longleftrightarrow P_{b_1\|a_1}$	$I_{a_1} \wedge I_{b_2} \Longleftrightarrow P_{b_2\|a_1}$
	$I_{a_2} \wedge I_{b_1} \Longleftrightarrow P_{b_1\|a_2}$	$I_{a_2} \wedge I_{b_2} \Longleftrightarrow P_{b_2\|a_2}$
C	$I_{a_1} \wedge I_{c_1} \Longleftrightarrow P_{c_1\|a_1}$	$I_{a_1} \wedge I_{c_2} \Longleftrightarrow P_{c_2\|a_1}$
	$I_{a_2} \wedge I_{c_1} \Longleftrightarrow P_{c_1\|a_2}$	$I_{a_2} \wedge I_{c_2} \Longleftrightarrow P_{c_2\|a_2}$

Encoding Local Structure

$$I_a \vee I_{\bar a} \qquad I_b \vee I_{\bar b} \qquad I_c \vee I_{\bar c}$$

$$\neg I_a \vee \neg I_{\bar a} \qquad \neg I_b \vee \neg I_{\bar b} \qquad \neg I_c \vee \neg I_{\bar c}$$

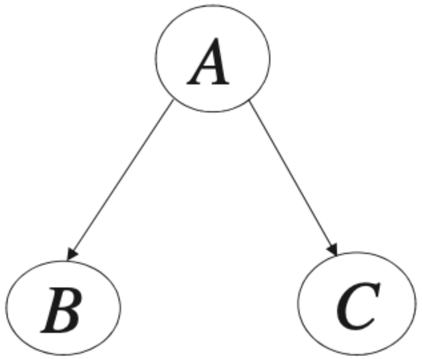
Zero Para

$$I_a \wedge I_{\bar b} \iff P_{\bar b|a}$$

$$I_a \iff P_a$$
$$I_{\bar a} \iff P_{\bar a}$$

$$I_a \wedge I_b \iff P_{b|a}$$
$$I_a \wedge I_{\bar b} \iff P_{\bar b|a}$$
$$I_{\bar a} \wedge I_b \iff P_{b|\bar a}$$
$$I_{\bar a} \wedge I_{\bar b} \iff P_{\bar b|\bar a}$$

$$I_a \wedge I_c \iff P_{c|a}$$
$$I_a \wedge I_{\bar c} \iff P_{\bar c|a}$$
$$I_{\bar a} \wedge I_c \iff P_{c|\bar a}$$
$$I_{\bar a} \wedge I_{\bar c} \iff P_{\bar c|\bar a}$$



A	θ_A
true	.5
false	.5

A	B	$\theta_{B\mid A}$
true	true	1
true	false	0
false	true	0
false	false	1

A	C	$\theta_{C\mid A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8

Encoding Local Structure

$$I_a \vee I_{\bar a} \qquad I_b \vee I_{\bar b} \qquad I_c \vee I_{\bar c}$$

$$\neg I_a \vee \neg I_{\bar a} \qquad \neg I_b \vee \neg I_{\bar b} \qquad \neg I_c \vee \neg I_{\bar c}$$

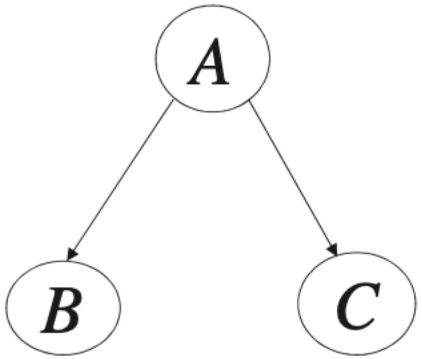
Zero Para

$$\neg I_a \vee \neg I_{\bar b}$$

$$I_a \iff P_a$$
$$I_{\bar a} \iff P_{\bar a}$$

$$I_a \wedge I_b \iff P_{b|a}$$
$$I_a \wedge I_{\bar b} \iff P_{\bar b|a}$$
$$I_{\bar a} \wedge I_b \iff P_{b|\bar a}$$
$$I_{\bar a} \wedge I_{\bar b} \iff P_{\bar b|\bar a}$$

$$I_a \wedge I_c \iff P_{c|a}$$
$$I_a \wedge I_{\bar c} \iff P_{\bar c|a}$$
$$I_{\bar a} \wedge I_c \iff P_{c|\bar a}$$
$$I_{\bar a} \wedge I_{\bar c} \iff P_{\bar c|\bar a}$$



A	θ_A
true	.5
false	.5

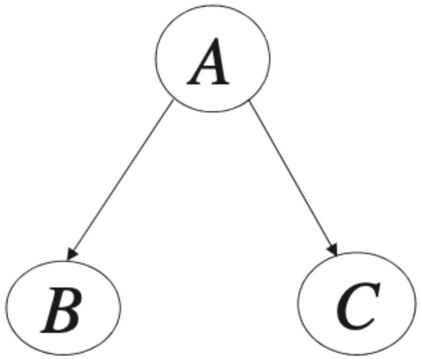
A	B	$\theta_{B\|A}$
true	true	1
true	false	0
false	true	0
false	false	1

A	C	$\theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8

Encoding Local Structure

One Para

$$I_a \wedge I_b \iff P_{b|a}$$



A	θ_A
true	.5
false	.5

A	B	$\theta_{B\|A}$
true	true	1
true	false	0
false	true	0
false	false	1

A	C	$\theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8

$$I_a \vee I_{\bar{a}} \qquad I_b \vee I_{\bar{b}} \qquad I_c \vee I_{\bar{c}}$$

$$\neg I_a \vee \neg I_{\bar{a}} \qquad \neg I_b \vee \neg I_{\bar{b}} \qquad \neg I_c \vee \neg I_{\bar{c}}$$

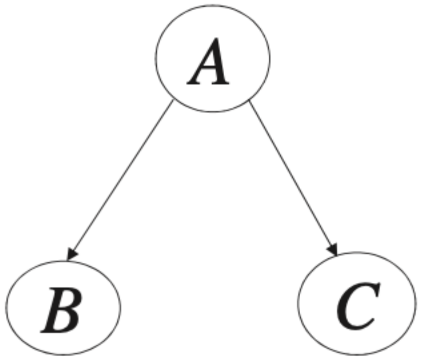
$$\begin{aligned}
I_a &\iff P_a \\
I_{\bar{a}} &\iff P_{\bar{a}}
\end{aligned}$$

$$\begin{aligned}
I_a \wedge I_b &\iff P_{b|a} \\
I_a \wedge I_{\bar{b}} &\iff P_{\bar{b}|a} \\
I_{\bar{a}} \wedge I_b &\iff P_{b|\bar{a}} \\
I_{\bar{a}} \wedge I_{\bar{b}} &\iff P_{\bar{b}|\bar{a}}
\end{aligned}$$

$$\begin{aligned}
I_a \wedge I_c &\iff P_{c|a} \\
I_a \wedge I_{\bar{c}} &\iff P_{\bar{c}|a} \\
I_{\bar{a}} \wedge I_c &\iff P_{c|\bar{a}} \\
I_{\bar{a}} \wedge I_{\bar{c}} &\iff P_{\bar{c}|\bar{a}}
\end{aligned}$$

Encoding Local Structure

One Para



A	θ_A
true	.5
false	.5

A	B	$\theta_{B\|A}$
true	true	1
true	false	0
false	true	0
false	false	1

A	C	$\theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8

$$I_a \vee I_{\bar{a}} \qquad I_b \vee I_{\bar{b}} \qquad I_c \vee I_{\bar{c}}$$

$$\neg I_a \vee \neg I_{\bar{a}} \qquad \neg I_b \vee \neg I_{\bar{b}} \qquad \neg I_c \vee \neg I_{\bar{c}}$$

$$\begin{aligned} I_a &\iff P_a \\ I_{\bar{a}} &\iff P_{\bar{a}} \end{aligned}$$

$$\begin{aligned} I_a \wedge I_b &\iff P_{b|a} \\ I_a \wedge I_{\bar{b}} &\iff P_{\bar{b}|a} \\ I_{\bar{a}} \wedge I_b &\iff P_{b|\bar{a}} \\ I_{\bar{a}} \wedge I_{\bar{b}} &\iff P_{\bar{b}|\bar{a}} \end{aligned}$$

$$\begin{aligned} I_a \wedge I_c &\iff P_{c|a} \\ I_a \wedge I_{\bar{c}} &\iff P_{\bar{c}|a} \\ I_{\bar{a}} \wedge I_c &\iff P_{c|\bar{a}} \\ I_{\bar{a}} \wedge I_{\bar{c}} &\iff P_{\bar{c}|\bar{a}} \end{aligned}$$

Encoding Local Structure

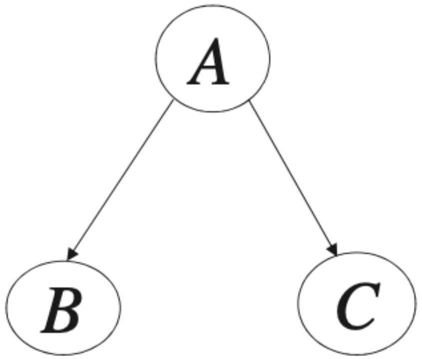
$$I_a \vee I_{\bar a} \qquad I_b \vee I_{\bar b} \qquad I_c \vee I_{\bar c}$$

$$\neg I_a \vee \neg I_{\bar a} \qquad \neg I_b \vee \neg I_{\bar b} \qquad \neg I_c \vee \neg I_{\bar c}$$

Equal Par

$$I_a \wedge I_c \iff P_{c|a}$$

$$I_{\bar a} \wedge I_{\bar c} \iff P_{\bar c|\bar a}$$



A	θ_A
true	.5
false	.5

A	B	$\theta_{B\|A}$
true	true	1
true	false	0
false	true	0
false	false	1

A	C	$\theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8

$$I_a \iff P_a$$
$$I_{\bar a} \iff P_{\bar a}$$

$$I_a \wedge I_b \iff P_{b|a}$$
$$I_a \wedge I_{\bar b} \iff P_{\bar b|a}$$
$$I_{\bar a} \wedge I_b \iff P_{b|\bar a}$$
$$I_{\bar a} \wedge I_{\bar b} \iff P_{\bar b|\bar a}$$

$$I_a \wedge I_c \iff P_{c|a}$$
$$I_a \wedge I_{\bar c} \iff P_{\bar c|a}$$
$$I_{\bar a} \wedge I_c \iff P_{c|\bar a}$$
$$I_{\bar a} \wedge I_{\bar c} \iff P_{\bar c|\bar a}$$

Encoding Local Structure

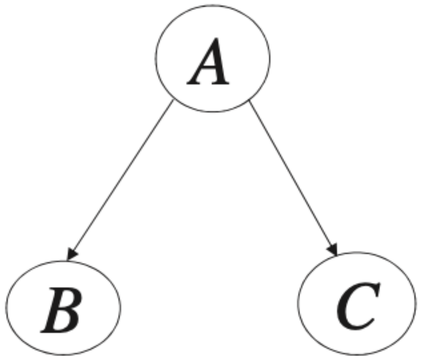
$$I_a \vee I_{\bar a} \qquad I_b \vee I_{\bar b} \qquad I_c \vee I_{\bar c}$$

$$\neg I_a \vee \neg I_{\bar a} \qquad \neg I_b \vee \neg I_{\bar b} \qquad \neg I_c \vee \neg I_{\bar c}$$

Equal Par

$$(I_a \wedge I_c) \vee (I_{\bar a} \wedge I_{\bar c}) \iff P_1$$

$$
\begin{aligned}
I_a &\iff P_a \\
I_{\bar a} &\iff P_{\bar a}
\end{aligned}
$$



$$
\begin{aligned}
I_a \wedge I_b &\iff P_{b|a} \\
I_a \wedge I_{\bar b} &\iff P_{\bar b|a} \\
I_{\bar a} \wedge I_b &\iff P_{b|\bar a} \\
I_{\bar a} \wedge I_{\bar b} &\iff P_{\bar b|\bar a}
\end{aligned}
$$

A	θ_A
true	.5
false	.5

A	B	$\theta_{B\|A}$
true	true	1
true	false	0
false	true	0
false	false	1

A	C	$\theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8

$$
\begin{aligned}
I_a \wedge I_c &\iff P_{c|a} \\
I_a \wedge I_{\bar c} &\iff P_{\bar c|a} \\
I_{\bar a} \wedge I_c &\iff P_{c|\bar a} \\
I_{\bar a} \wedge I_{\bar c} &\iff P_{\bar c|\bar a}
\end{aligned}
$$

Encoding Local Structure

A	θ_A
true	.5
false	.5

A	B	$\theta_{B\mid A}$
true	true	1
true	false	0
false	true	0
false	false	1

A	C	$\theta_{C\mid A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8

without local structure

$$I_a \vee I_{\bar{a}} \qquad I_b \vee I_{\bar{b}} \qquad I_c \vee I_{\bar{c}}$$
$$\neg I_a \vee \neg I_{\bar{a}} \qquad \neg I_b \vee \neg I_{\bar{b}} \qquad \neg I_c \vee \neg I_{\bar{c}}$$

$$\begin{aligned}
I_a &\iff P_a \\
I_{\bar{a}} &\iff P_{\bar{a}}
\end{aligned}$$

$$\begin{aligned}
I_a \wedge I_b &\iff P_{b|a} \\
I_a \wedge I_{\bar{b}} &\iff P_{\bar{b}|a} \\
I_{\bar{a}} \wedge I_b &\iff P_{b|\bar{a}} \\
I_{\bar{a}} \wedge I_{\bar{b}} &\iff P_{\bar{b}|\bar{a}}
\end{aligned}$$

$$\begin{aligned}
I_a \wedge I_c &\iff P_{c|a} \\
I_a \wedge I_{\bar{c}} &\iff P_{\bar{c}|a} \\
I_{\bar{a}} \wedge I_c &\iff P_{c|\bar{a}} \\
I_{\bar{a}} \wedge I_{\bar{c}} &\iff P_{\bar{c}|\bar{a}}
\end{aligned}$$

with local structure

$$I_a \vee I_{\bar{a}} \qquad I_b \vee I_{\bar{b}} \qquad I_c \vee I_{\bar{c}}$$
$$\neg I_a \vee \neg I_{\bar{a}} \qquad \neg I_b \vee \neg I_{\bar{b}} \qquad \neg I_c \vee \neg I_{\bar{c}}$$

$$\neg I_a \vee \neg I_{\bar{b}}$$
$$\neg I_{\bar{a}} \vee \neg I_b$$

$$\begin{aligned}
(I_a \wedge I_c) \vee (I_{\bar{a}} \wedge I_{\bar{c}}) &\iff P_1 \quad \text{(weight of } P_1 \text{ is .8)} \\
(I_a \wedge I_{\bar{c}}) \vee (I_{\bar{a}} \wedge I_c) &\iff P_2 \quad \text{(weight of } P_2 \text{ is .2)} \\
I_a \vee I_{\bar{a}} &\iff P_3 \quad \text{(weight of } P_3 \text{ is .5)}
\end{aligned}$$

WMC using Tractable Circuits

$$I_a \vee I_{\bar a} \qquad I_b \vee I_{\bar b} \qquad I_c \vee I_{\bar c}$$

$$\neg I_a \vee \neg I_{\bar a} \qquad \neg I_b \vee \neg I_{\bar b} \qquad \neg I_c \vee \neg I_{\bar c}$$

$$\neg I_a \vee \neg I_{\bar b}$$

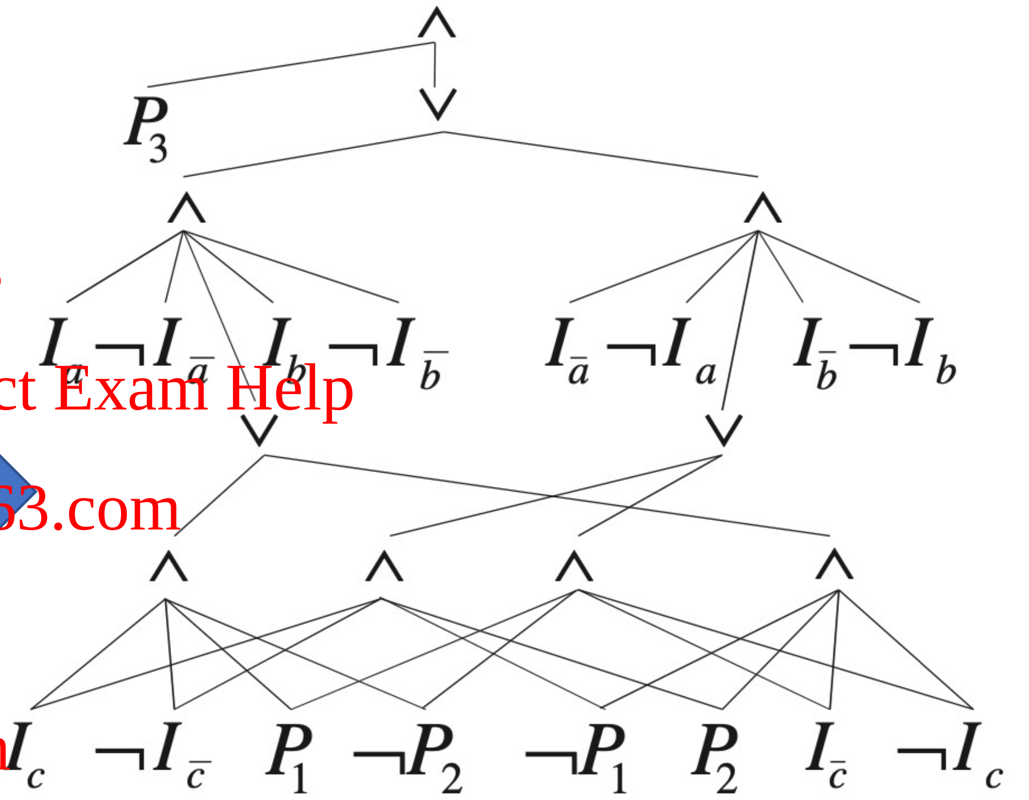
$$\neg I_{\bar a} \vee \neg I_b$$

$$(I_a \wedge I_c) \vee (I_{\bar a} \wedge I_{\bar c}) \iff P_1 \quad \text{(weight of } P_1 \text{ is .8)}$$

$$(I_a \wedge I_{\bar c}) \vee (I_{\bar a} \wedge I_c) \iff P_2 \quad \text{(weight of } P_2 \text{ is .2)}$$

$$I_a \vee I_{\bar a} \iff P_3 \quad \text{(weight of } P_3 \text{ is .5)}$$

compile



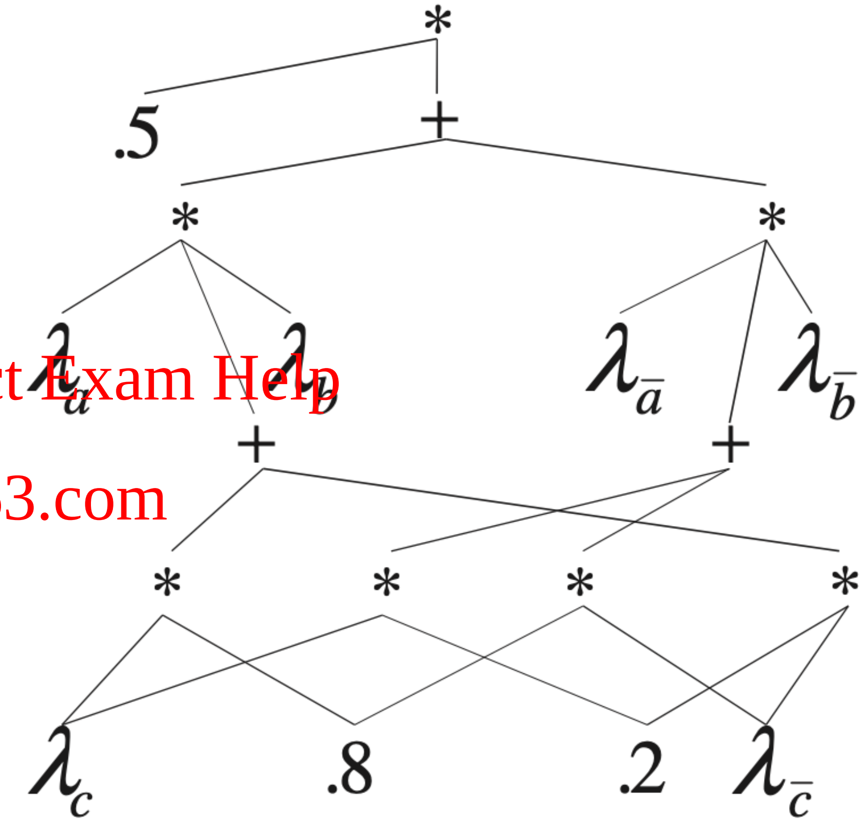
NNF Circuit: Decomposable, Deterministic, Smooth

and → *
or → +
negative literals → 1
positive parameter literals → weight

WMC using Tractable Circuits

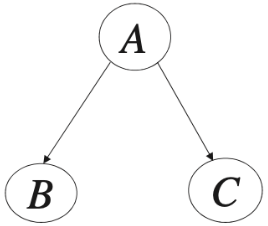
$I_a \vee I_{\bar{a}}$	$I_b \vee I_{\bar{b}}$	$I_c \vee I_{\bar{c}}$
$\neg I_a \vee \neg I_{\bar{a}}$	$\neg I_b \vee \neg I_{\bar{b}}$	$\neg I_c \vee \neg I_{\bar{c}}$
	$\neg I_a \vee \neg I_{\bar{b}}$	
	$\neg I_{\bar{a}} \vee \neg I_b$	

$$(I_a \wedge I_c) \vee (I_{\bar{a}} \wedge I_{\bar{c}}) \iff P_1 \quad \text{(weight of } P_1 \text{ is .8)}$$
$$(I_a \wedge I_{\bar{c}}) \vee (I_{\bar{a}} \wedge I_c) \iff P_2 \quad \text{(weight of } P_2 \text{ is .2)}$$
$$I_a \vee I_{\bar{a}} \iff P_3 \quad \text{(weight of } P_3 \text{ is .5)}$$



Arithmetic Circuit (AC)

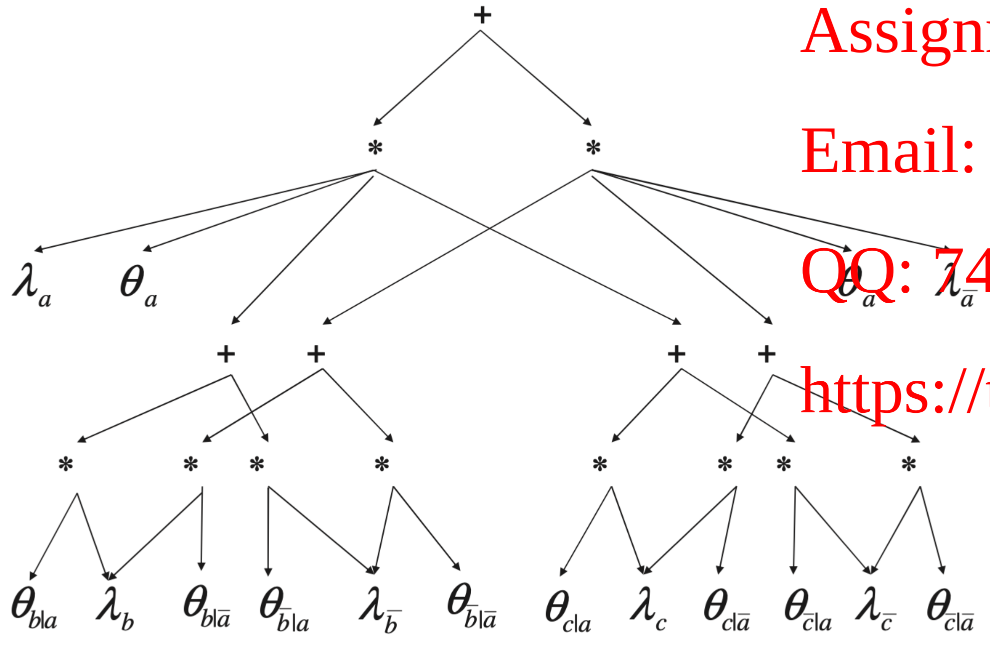
Local Structure



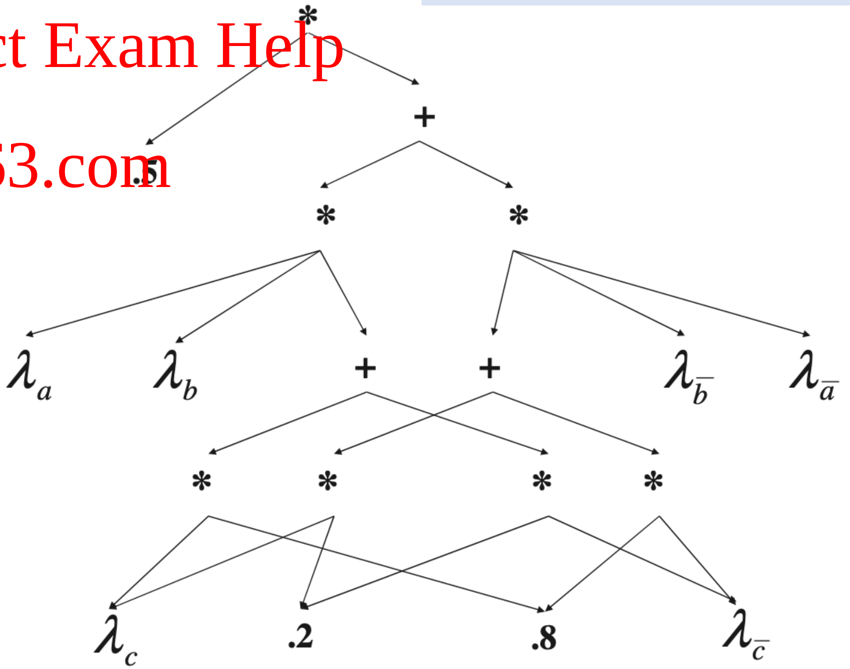
A	B	$\theta_{B\|A}$
true	true	1
true	false	0
false	true	0
false	false	1

A	C	$\theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8

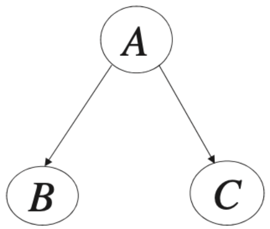
AC without local structure



AC with local structure

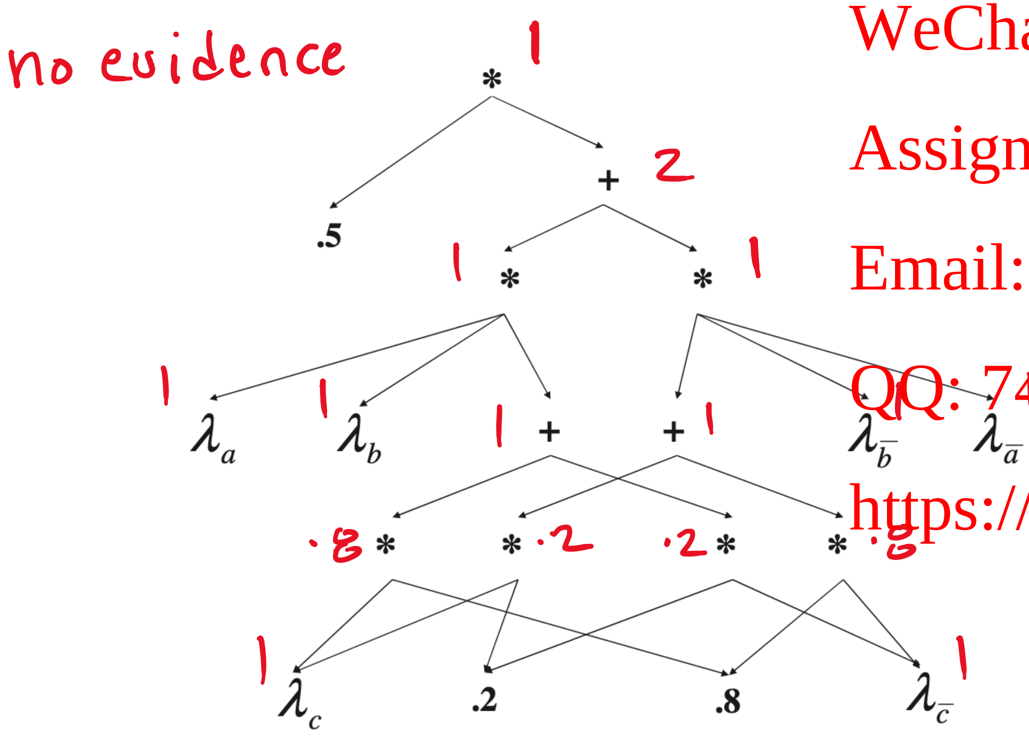


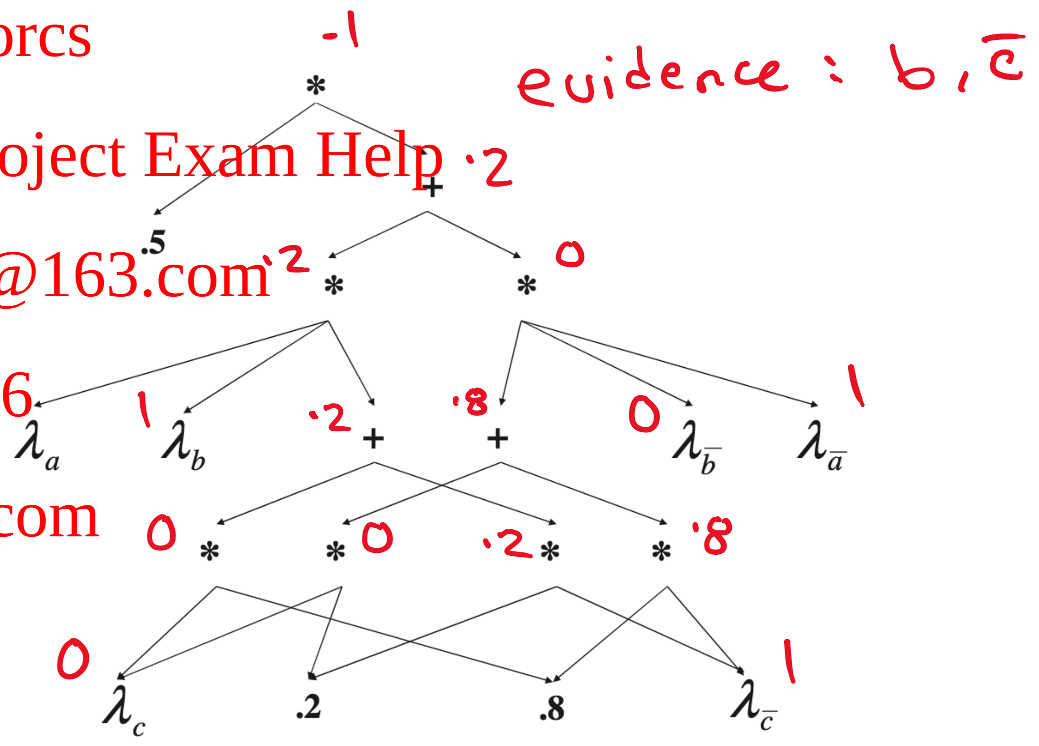
Reasoning with ACs



A	B	$\theta_{B\|A}$
true	true	1
true	false	0
false	true	0
false	false	1

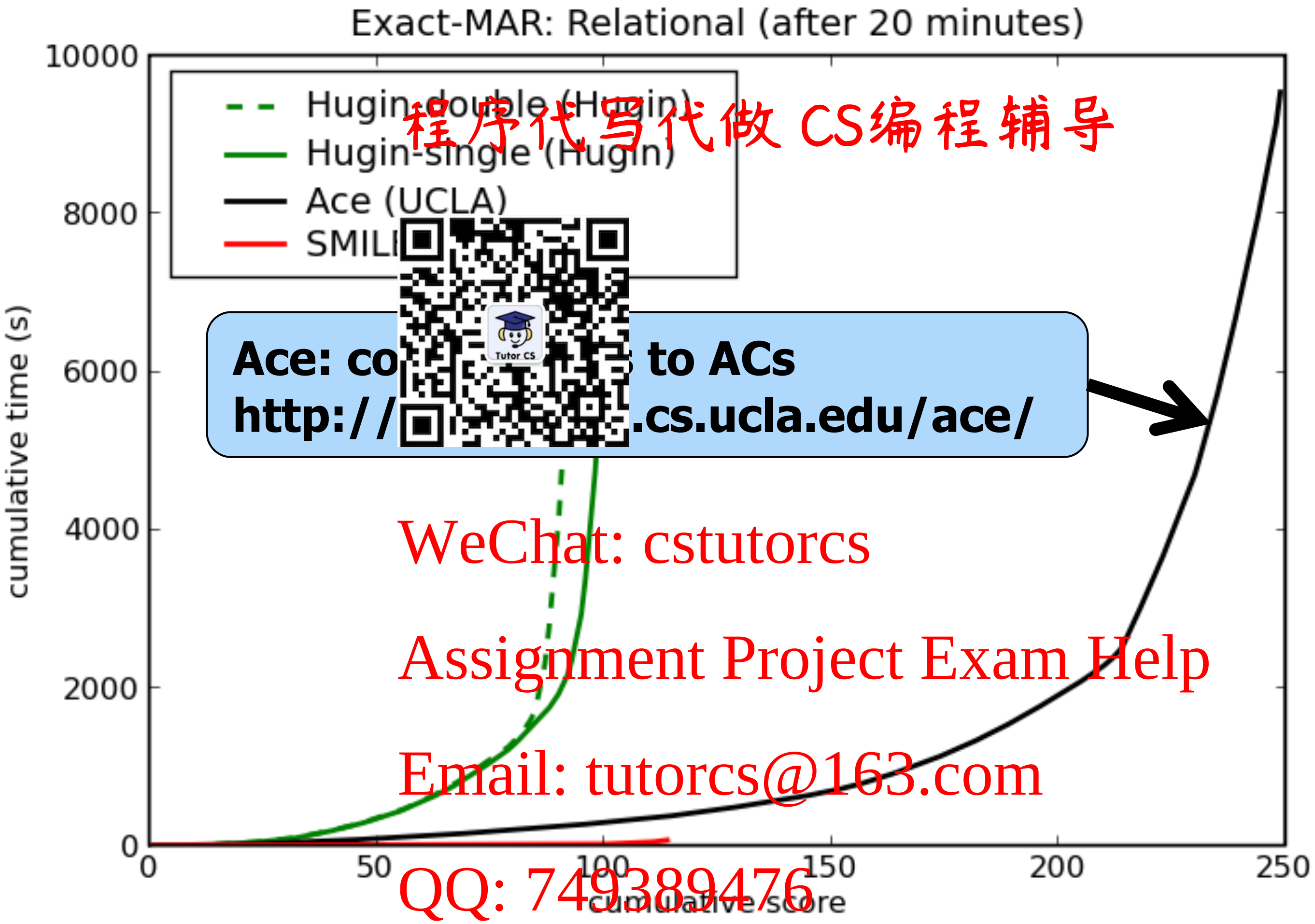
A	C	$\theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.2
false	false	.8





Classical Probabilistic Reasoning Algorithms

$$O(n\exp\{w\})$$



Relational networks (251 networks)

have massive logical constraints and context-specific independence

reductions now common in probabilistic programming

Other Reductions

- Variables with large cardinalities
- Better exploitation of local structure
- Minimize size of CNF
- ….

Durgesh Agrawal, Yash Pote, Kuldeep S. Meel:
Partition Function Estimation: A Quantitative Study. IJCAI-21

Method Name	Problem Classes								
	[illegible]	[illegible]ne- (65)	BN (60)	Ising (52)	Segment (50)	ObjDetect (35)	Protein (29)	Misc (27)	Total (672)
Ace	[illegible]	[illegible]	60	51	50	0	16	15	611
Fractional Belief Propagation (FBP)	293	65	58	41	48	32	29	9	575
Loopy Belief Propagation (BP)	292	65	58	41	46	32	29	10	573
Generalized Belief Propagation (GBP)	281	65	36	47	40	34	29	9	541
Edge Deletion Belief Propagation (EDBP)	245	42	56	50	49	35	28	23	528
GANAK	353	58	53	4	0	0	7	14	489
Double Loop Generalised BP (HAK)	199	65	58	43	43	35	29	14	486
Tree Expectation Propagation (TREEEP)	101	65	58	50	48	35	29	15	401
SampleSearch	89	56	33	52	37	35	29	25	356
Bucket Elimination (BE)	98	32	15	52	50	35	29	22	333
Conditioned Belief Propagation (CBP)	109	32	21	41	50	35	29	8	325
Join Tree (JT)	98	32	15	52	50	19	26	21	313
Dynamic Importance Sampling (DIS)	24	65	25	52	50	35	29	27	307
Weighted Mini Bucket Elimination (WMB)	68	13	17	50	50	20	28	12	258
miniC2D	187	1	30	31	0	0	0	1	250
WeightCount	93	0	27	0	0	0	0	0	120
WISH	0	0	0	9	0	0	0	0	9
FocusedFlatSAT	6	0	0	0	0	0	0	0	6

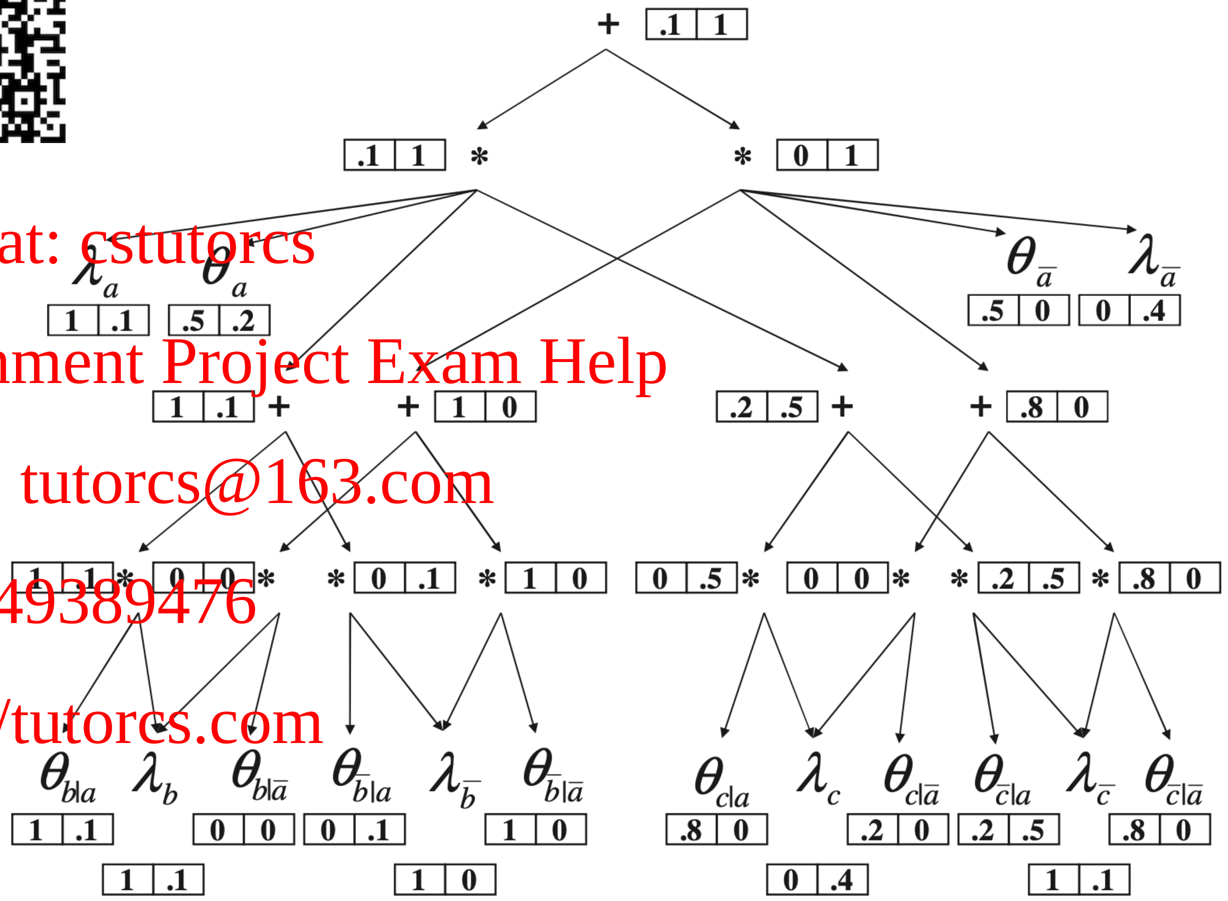
reasoning.cs.ucla.edu/ace

Backpropagation on ACs

Compute marginals in bu

$$\frac{\partial f}{\partial \lambda_x}(\mathbf{e}) = \Pr(x, \mathbf{e} - X).$$

$$\theta_{x|\mathbf{u}} \frac{\partial f}{\partial \theta_{x|\mathbf{u}}}(\mathbf{e}) = \Pr(x, \mathbf{u}, \mathbf{e}).$$

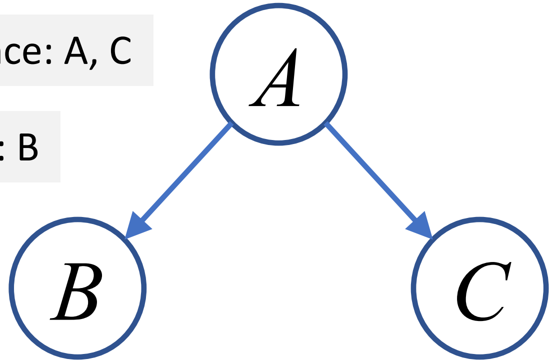


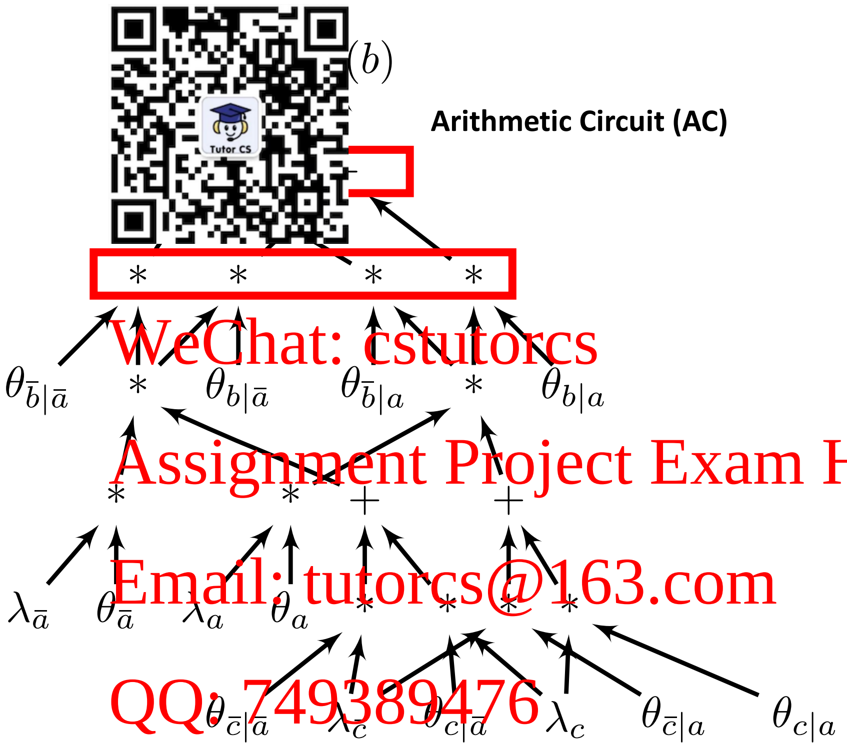
Further applications to learning parameters using gradient descent and EM

Model-Based Supervised Learning using ACs

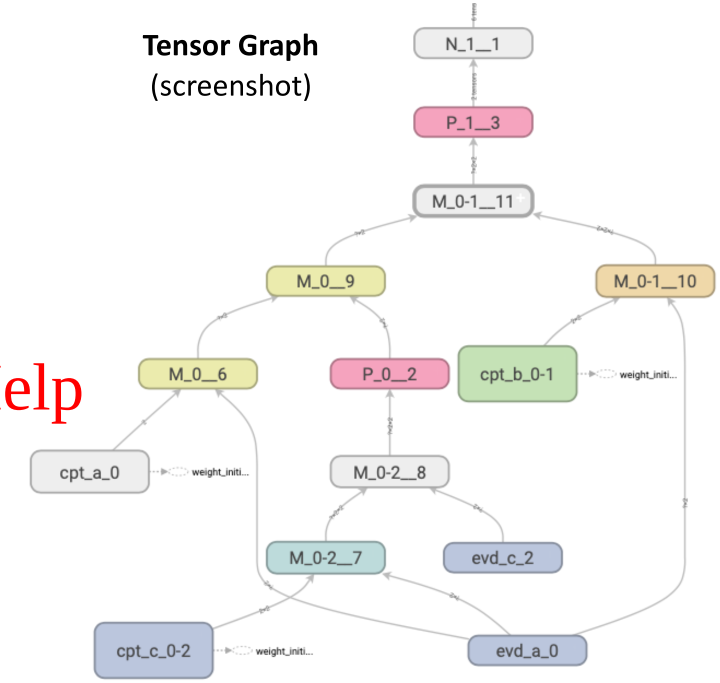
Evidence: A, C

Query: B





Tensor Graph (screenshot)



(mostly matrix multiplication)

can be trained discriminatively from labeled data using gradient descent

can integrate domain knowledge (independence, logical constraints, functional dependencies)

new AC compilation algorithm: can exploit functional dependencies computationally without knowing their identities

Arithmetic Circuits (ACs)

A Modern Perspective

From Boolean to Arithmetic Circuits

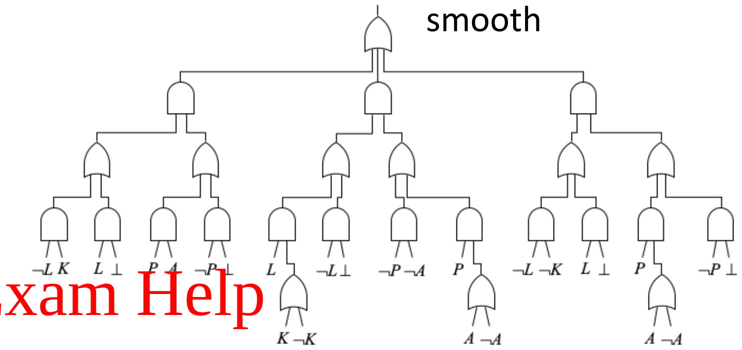
Bayesian Network

reduce

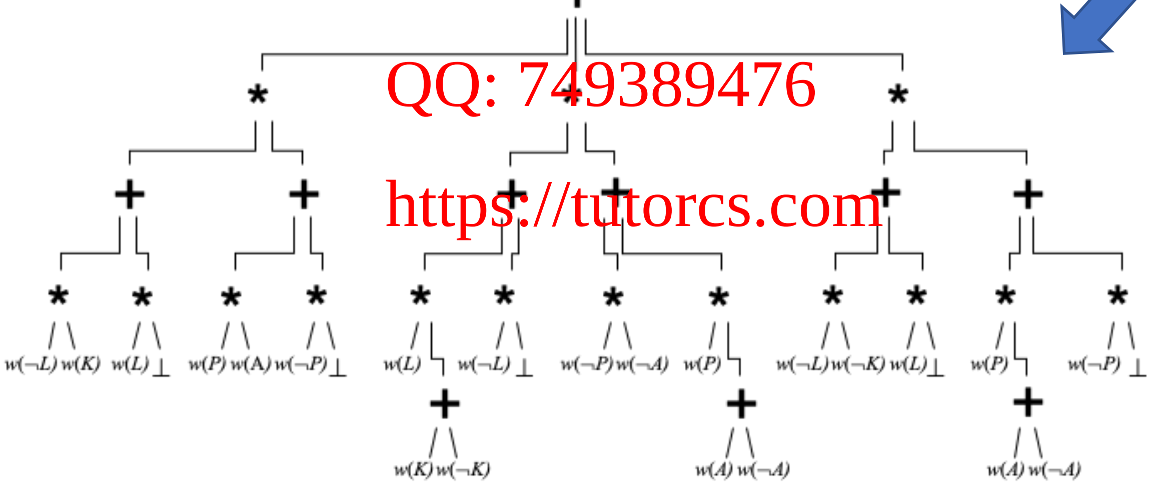
Weighted Boolean Formula

compile

decomposable
deterministic
smooth



Arithmetic Circuit (AC)



A Different Perspective

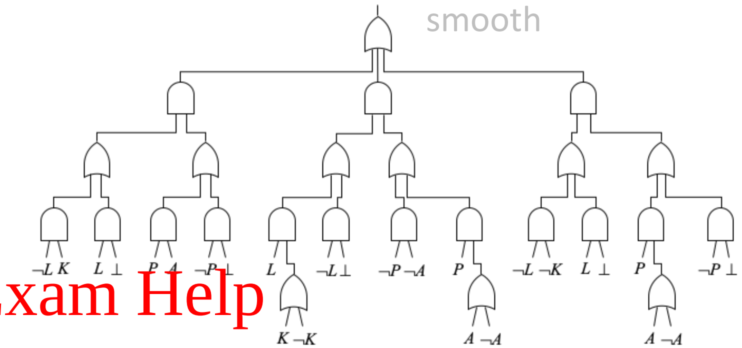
Bayesian Network

reduce

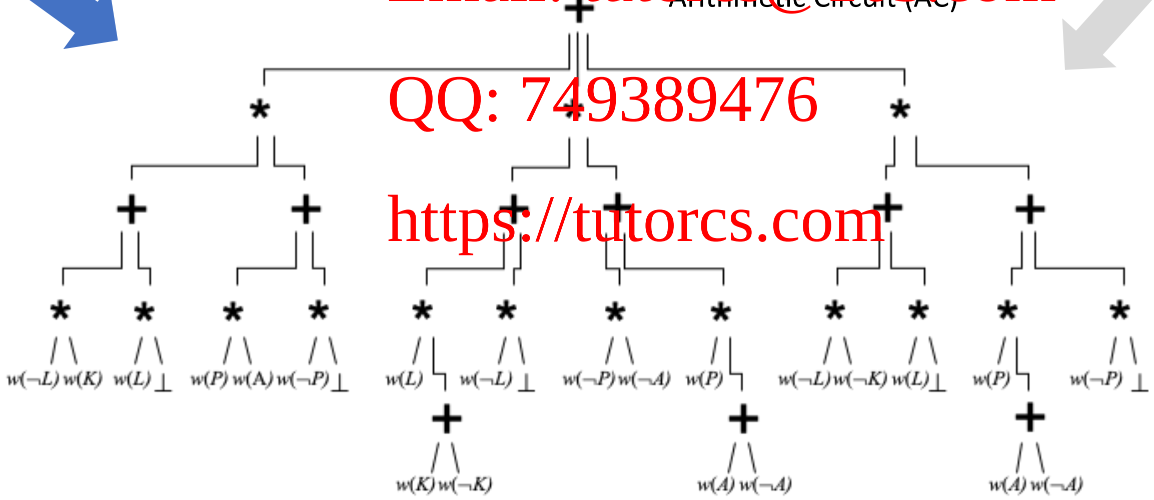
Weighted Boolean Formula

compile

decomposable
deterministic
smooth



Arithmetic Circuit (AC)



A Different Perspective

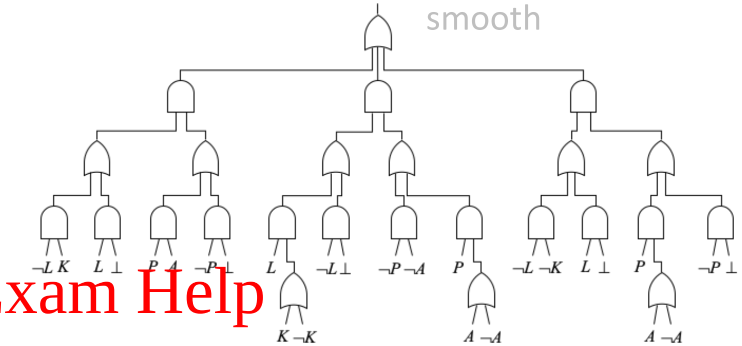
Factor

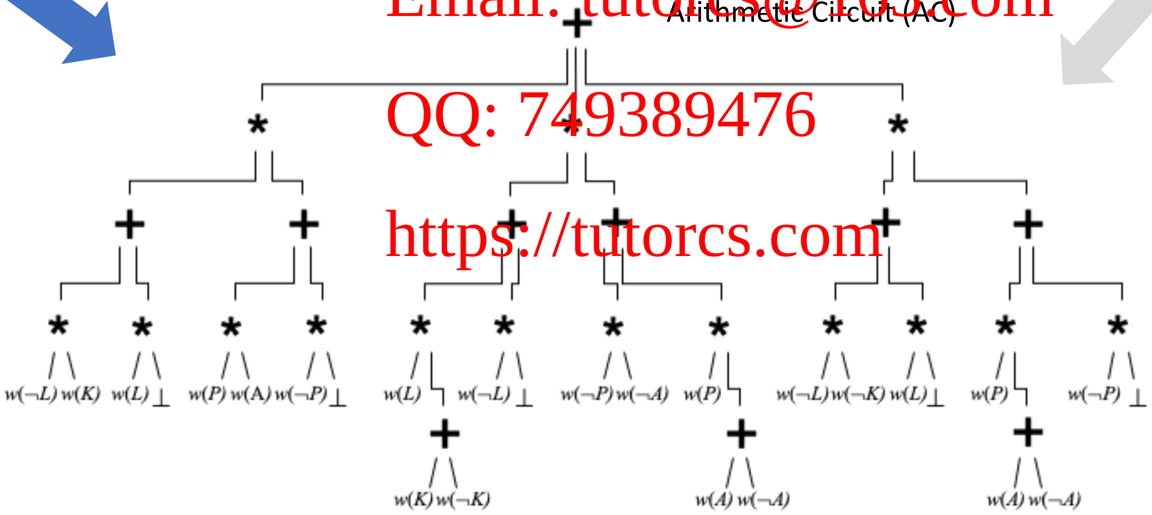
reduce

Weighted Boolean Formula

compile

decomposable
deterministic
smooth





Arithmetic Circuit (AC)

Learning ACs from data

Handcrafting ACs

Reference Distribution/Factor…

MPE and Marginals on Factors

Factor

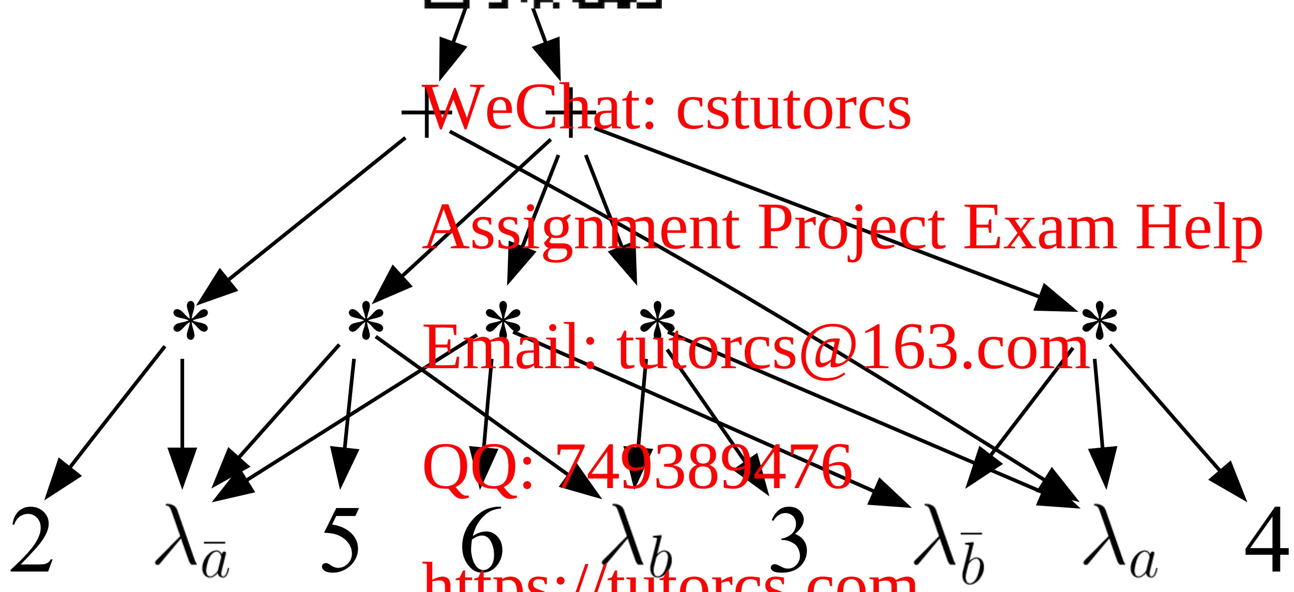
A	B	C	
T	T	T	0.08
T	T	F	1.24
T	F	T	0.10
T	F	F	3.10
F	T	T	0.00
F	T	F	0.00
F	F	T	9.42
F	F	F	0.06

MPE
find maximal value of row that satisfies some event (e.g., B=F)
(subsumes SAT)

MAR
sum values of rows that satisfy some condition (e.g., B=F)
(subsumes MAJ-SAT, #SAT and WMC)

Two Fundamental Notions
Three Fundamental Questions

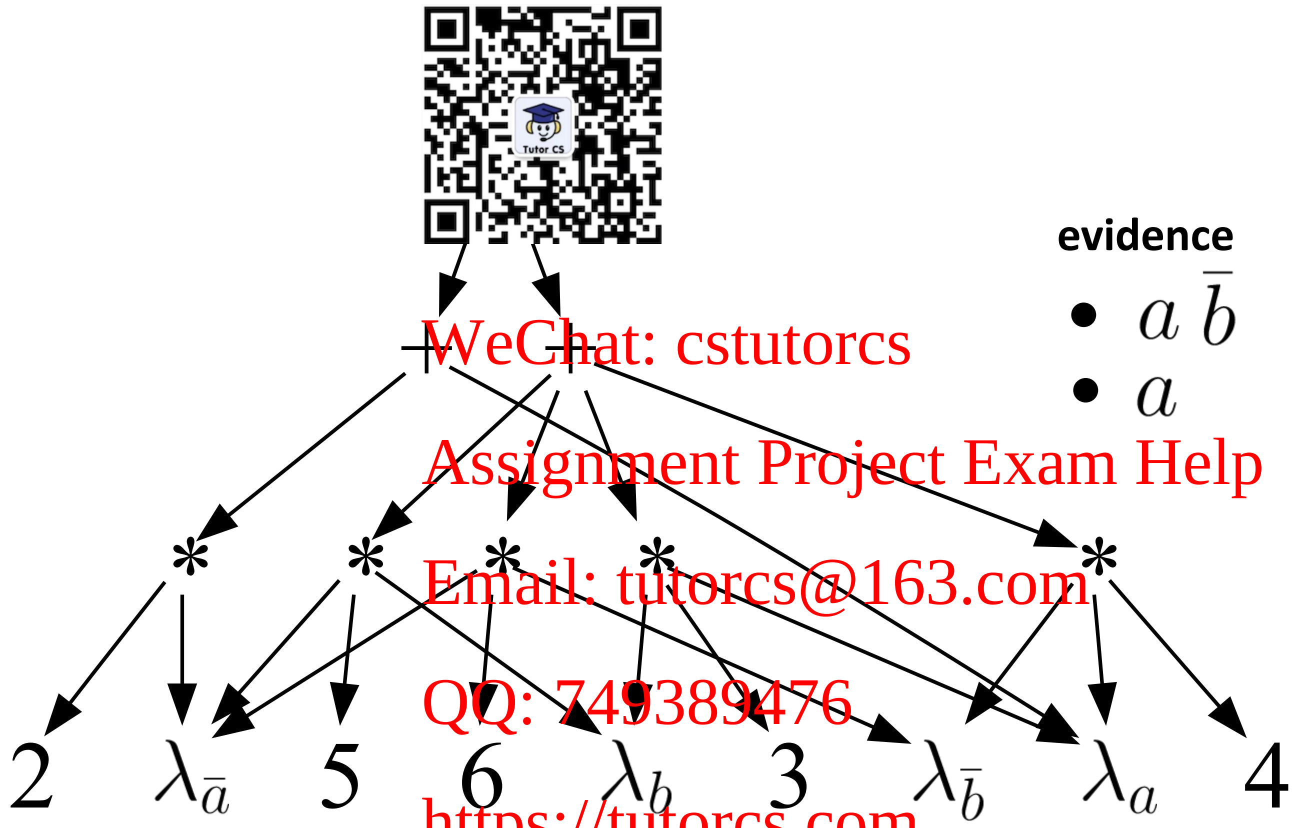
Fundamental Notion: Arithmetic Circuit (AC)



values of variables define indicators

indicator variables, constants, additions, multiplications

Fundamental Notion: Evaluating AC at evidence



evidence	λ_a	$\lambda_{\bar{a}}$	λ_b	$\lambda_{\bar{b}}$
	indicators			
• $a\,\bar{b}$	**1**	**0**	**0**	**1**
• a	**1**	**0**	**1**	**1**

indicator replaced by **1** if its subscript is **consistent** with **evidence**

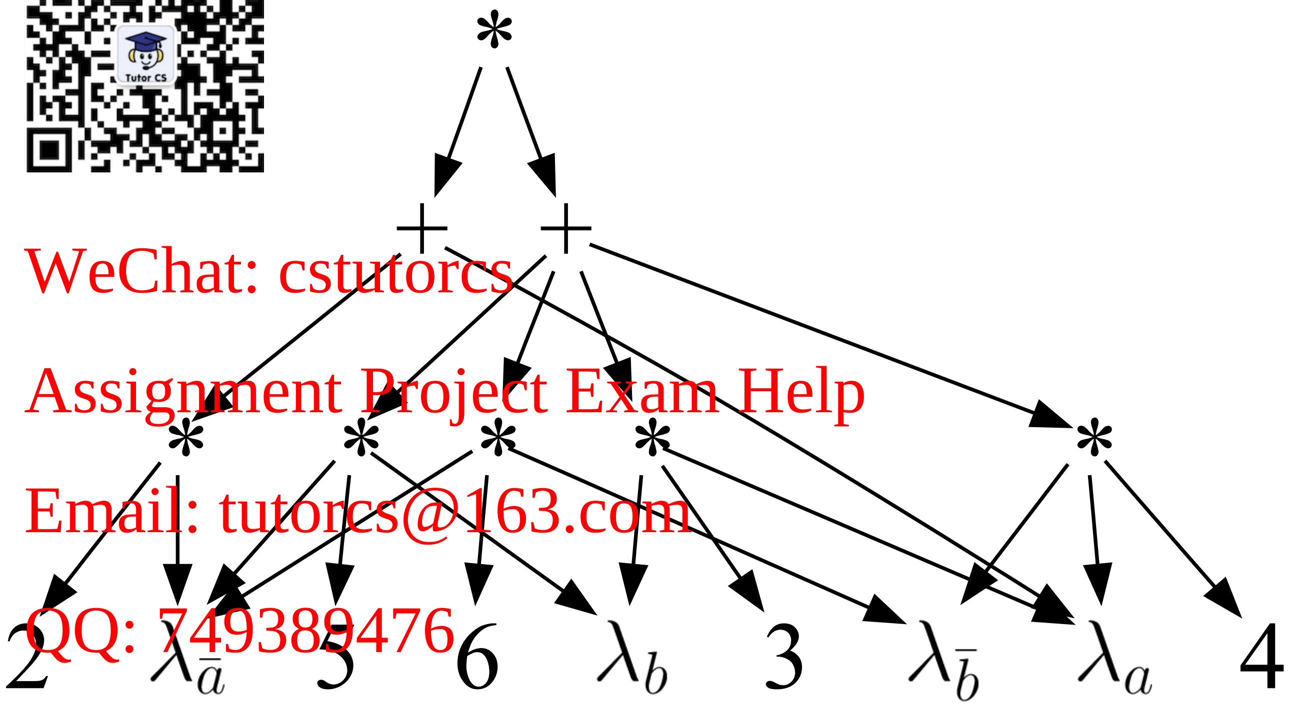
indicator replaced by **0** if its subscript is **inconsistent** with **evidence**

Three Fundamental Questions

- What factor f(**X**) does an AC **represent?**
  (the reference factor)
- When does an AC **compute marginals** of factor f(**X**)?
  (in linear time)
- When does an AC **compute MPE** of factor f(**X**)?
  (in linear time)

Factor Represented by an AC

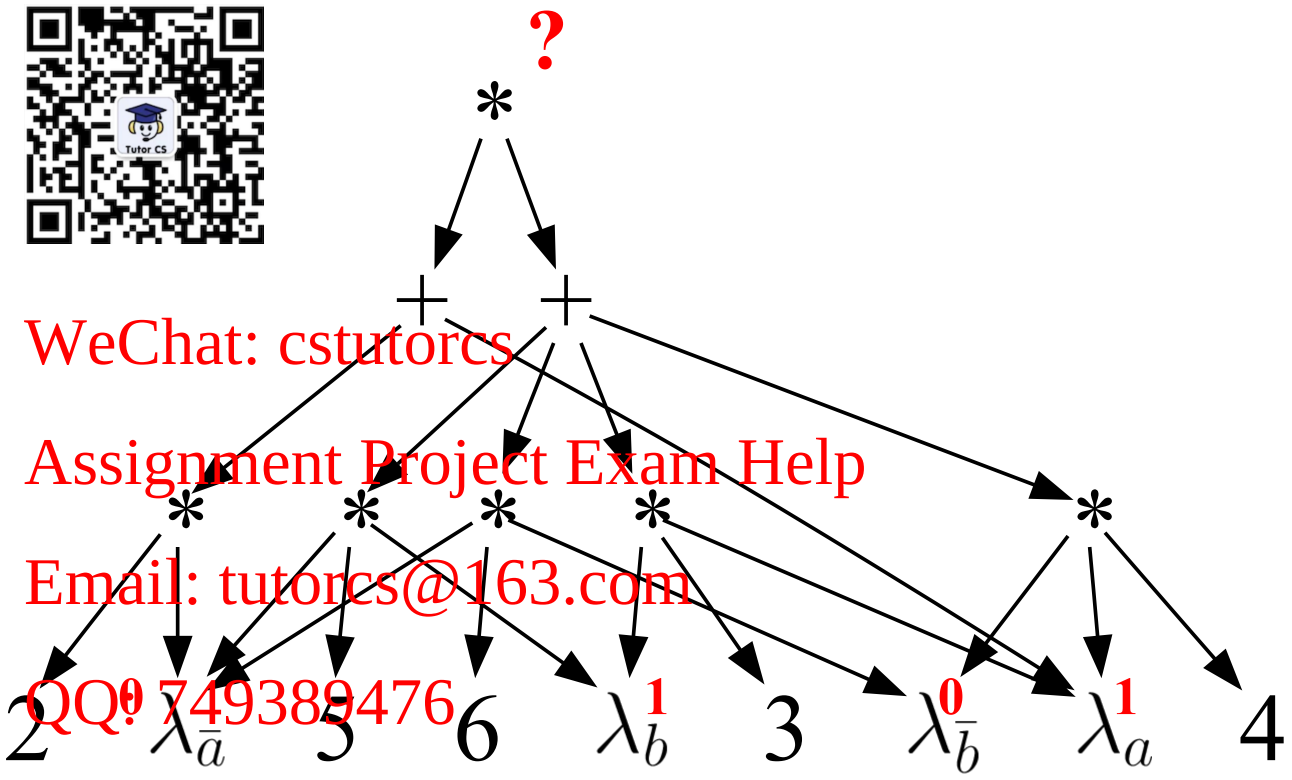
A	B	$f(A,B)$
a	b	?
a	$\bar{b}$	?
$\bar{a}$	b	?
$\bar{a}$	$\bar{b}$	?



evaluate AC at complete variable instantiations (evidence that fixes values of *A* and *B*)

Factor Represented by an AC

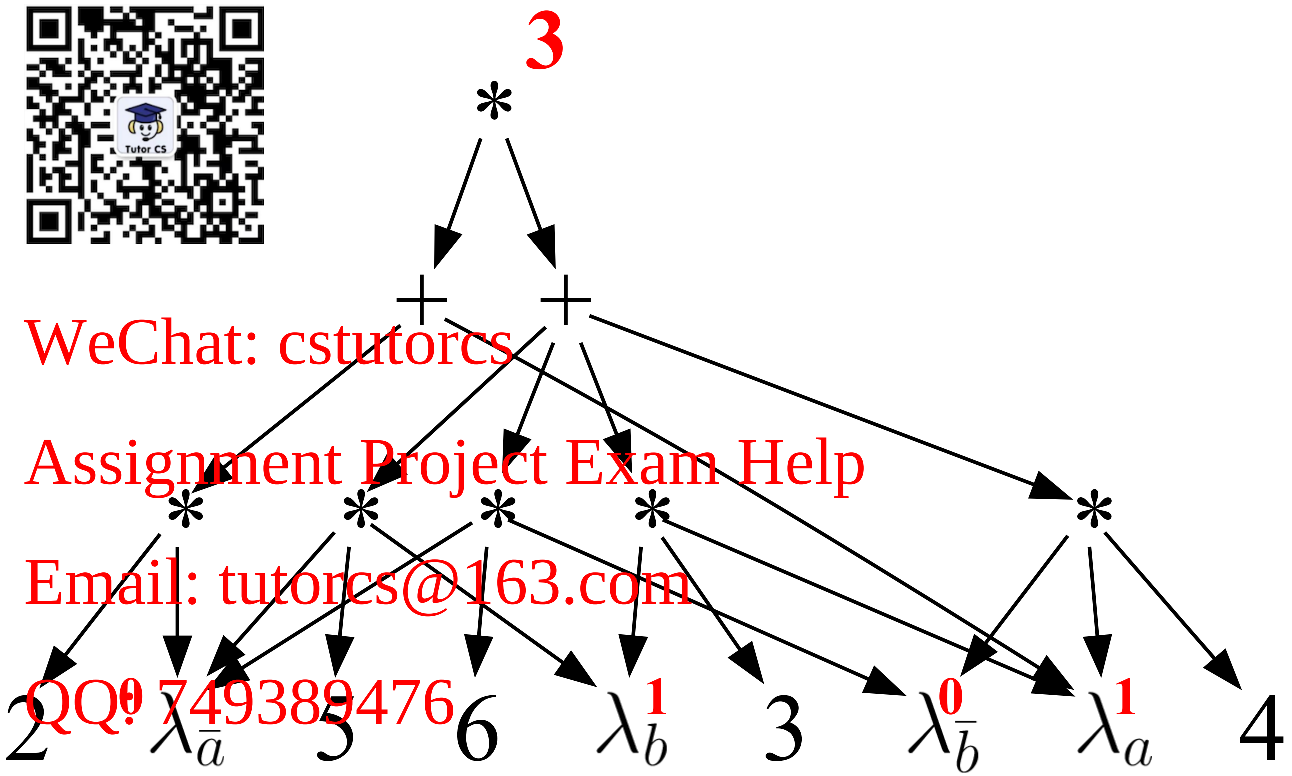
A	B	$f(A,B)$
a	b	**?**
a	$\bar{b}$	?
$\bar{a}$	b	?
$\bar{a}$	$\bar{b}$	?



evaluate AC at complete variable instantiations (evidence that fixes values of *A* and *B*)

Factor Represented by an AC

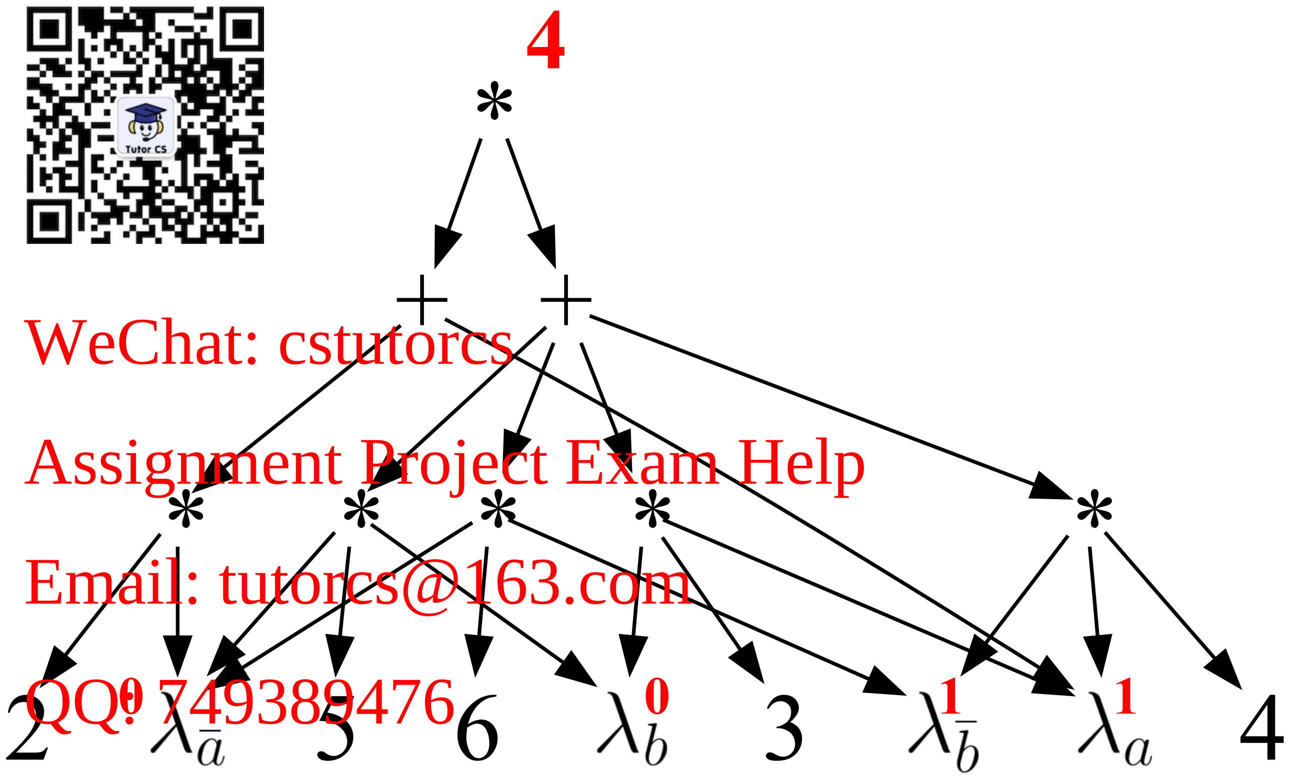
A	B	$f(A,B)$
a	b	**3**
a	$\bar{b}$	?
$\bar{a}$	b	?
$\bar{a}$	$\bar{b}$	?



evaluate AC at complete variable instantiations (evidence that fixes values of *A* and *B*)

Factor Represented by an AC

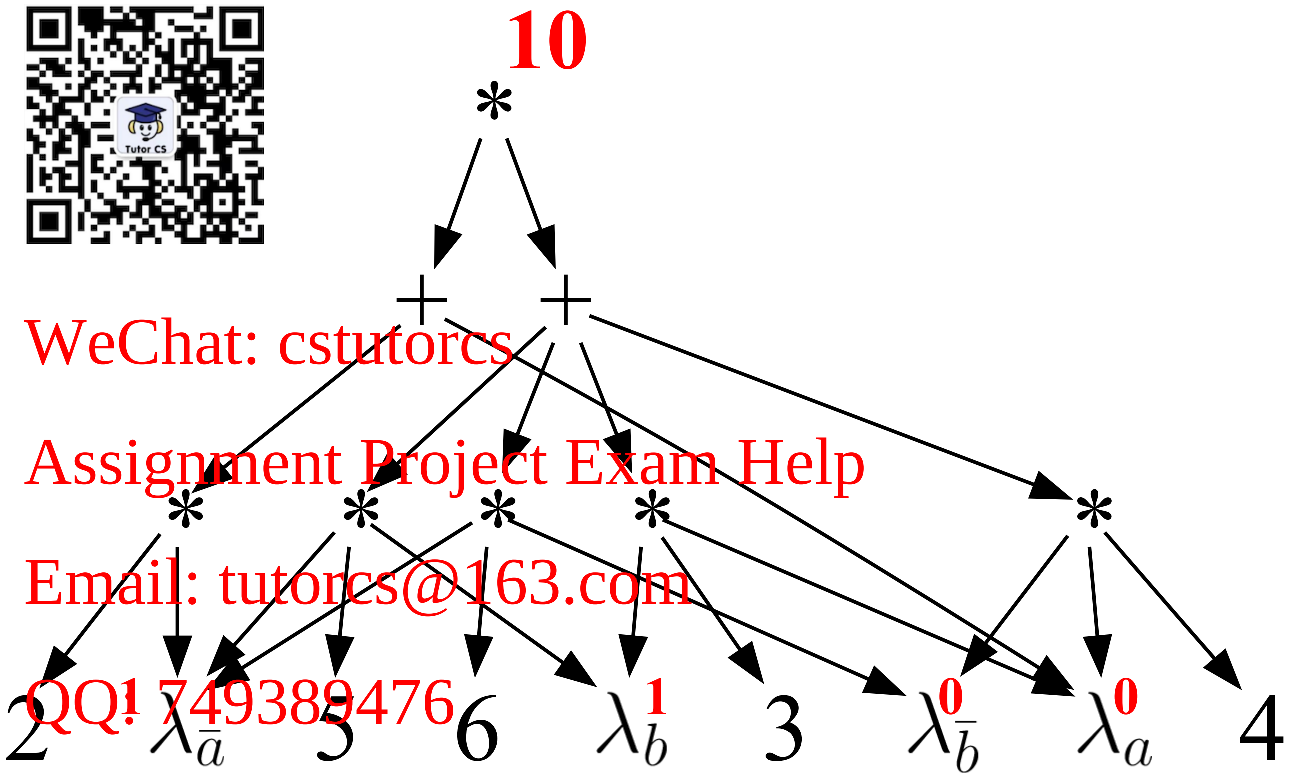
A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	**4**
$\bar{a}$	b	?
$\bar{a}$	$\bar{b}$	?



evaluate AC at complete variable instantiations (evidence that fixes values of *A* and *B*)

Factor Represented by an AC

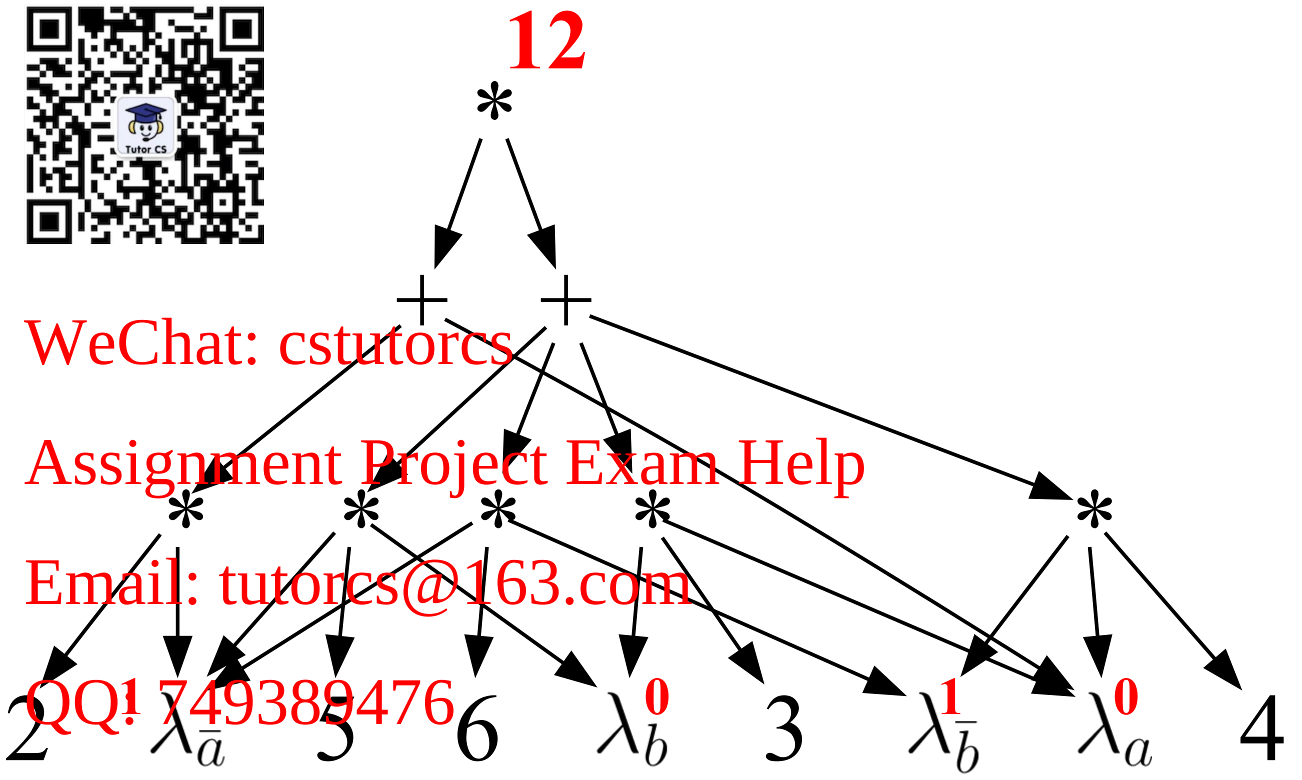
A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	**10**
$\bar{a}$	$\bar{b}$	?



evaluate AC at complete variable instantiations (evidence that fixes values of *A* and *B*)

Factor Represented by an AC

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	**12**



evaluate AC at complete variable instantiations (evidence that fixes values of *A* and *B*)

Key Observation

it is easy to construct an AC that represents the factor of a Bayesian network

A	$f_1(A)$
a	1
$\bar{a}$	2

A	B	$f_2(A, B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	5
$\bar{a}$	$\bar{b}$	6

A	B	$f(A, B)$
a	b	$1 * 3$
a	$\bar{b}$	$1 * 4$
$\bar{a}$	b	$2 * 5$
$\bar{a}$	$\bar{b}$	$2 * 6$

A	$f(A)$
a	$1 \cdot \lambda_a$
$\bar{a}$	$2 \cdot \lambda_{\bar{a}}$

A	B	$f(A, B)$
a	b	$3 \cdot \lambda_a \lambda_b$
a	$\bar{b}$	$4 \cdot \lambda_a \lambda_{\bar{b}}$
$\bar{a}$	b	$5 \cdot \lambda_{\bar{a}} \lambda_b$
$\bar{a}$	$\bar{b}$	$6 \cdot \lambda_{\bar{a}} \lambda_{\bar{b}}$

AC: Size linear in the size of input factors

$$[1 \cdot \lambda_a + 2 \cdot \lambda_{\bar{a}}] * [3 \cdot \lambda_a \lambda_b + 4 \cdot \lambda_a \lambda_{\bar{b}} + 5 \cdot \lambda_{\bar{a}} \lambda_b + 6 \cdot \lambda_{\bar{a}} \lambda_{\bar{b}}]$$

$$\begin{aligned} & AC(a, \bar{b}) \\ &= [1(1) + 2(0)] * [3(1)(0) + 4(1)(1) + 5(0)(0) + 6(0)(1)] \\ &= 1 * 4 \end{aligned}$$

$$1 \cdot \lambda_a + 2 \cdot \lambda_{\bar{a}}$$

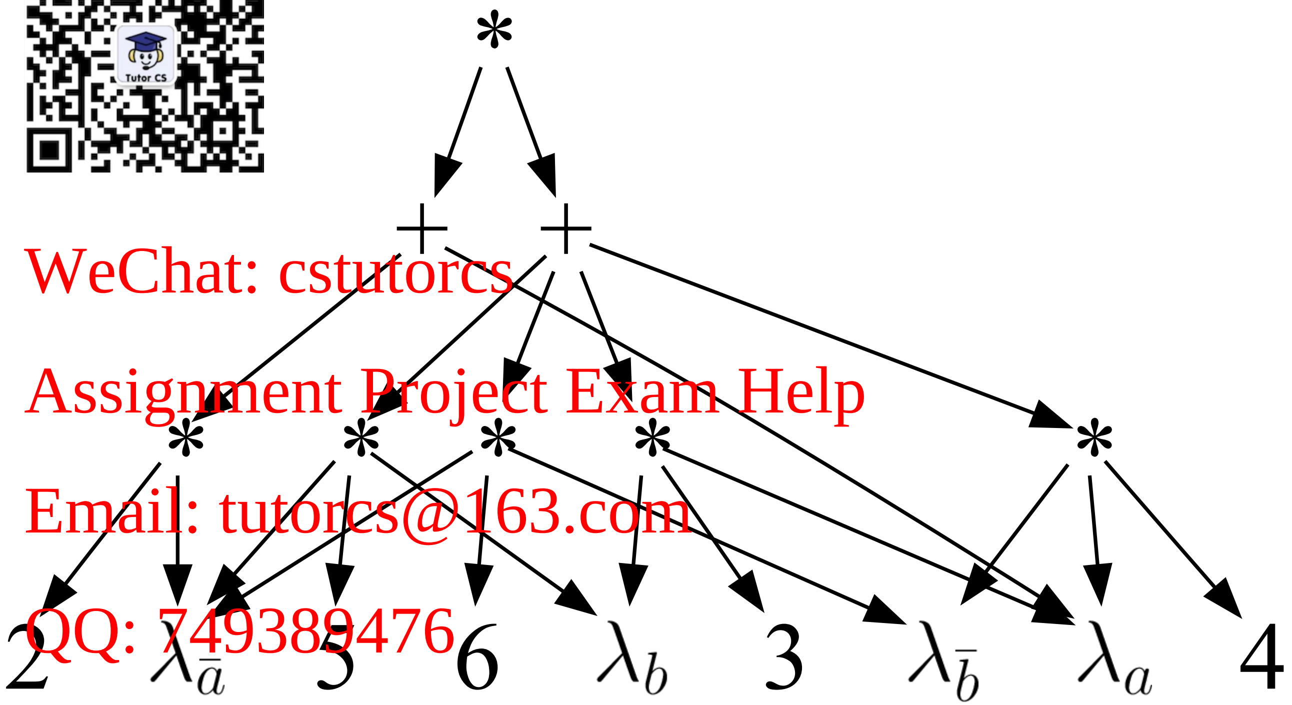
$$3 \cdot \lambda_a \lambda_b + 4 \cdot \lambda_a \lambda_{\bar{b}} + 5 \cdot \lambda_{\bar{a}} \lambda_b + 6 \cdot \lambda_{\bar{a}} \lambda_{\bar{b}}$$

Three Fundamental Questions

- What factor f(**X**) does an AC **represent?**
  (the reference factor)
- When does an AC **compute marginals** of factor f(**X**)?
  (in linear time)
- When does an AC **compute MPE** of factor f(**X**)?
  (in linear time)

Does AC Compute Marginals?

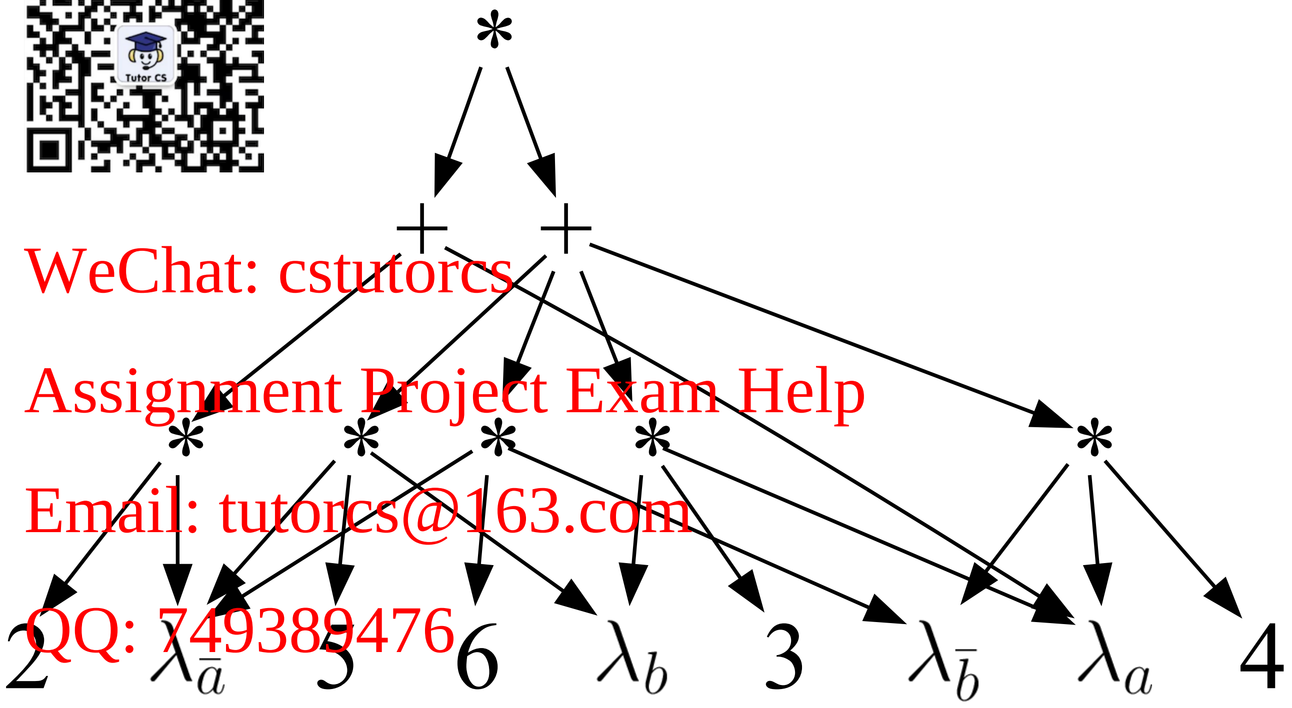
A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12



Does AC evaluate correctly at every partial variable instantiation (evidence)?

Does AC Compute Marginals?

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12

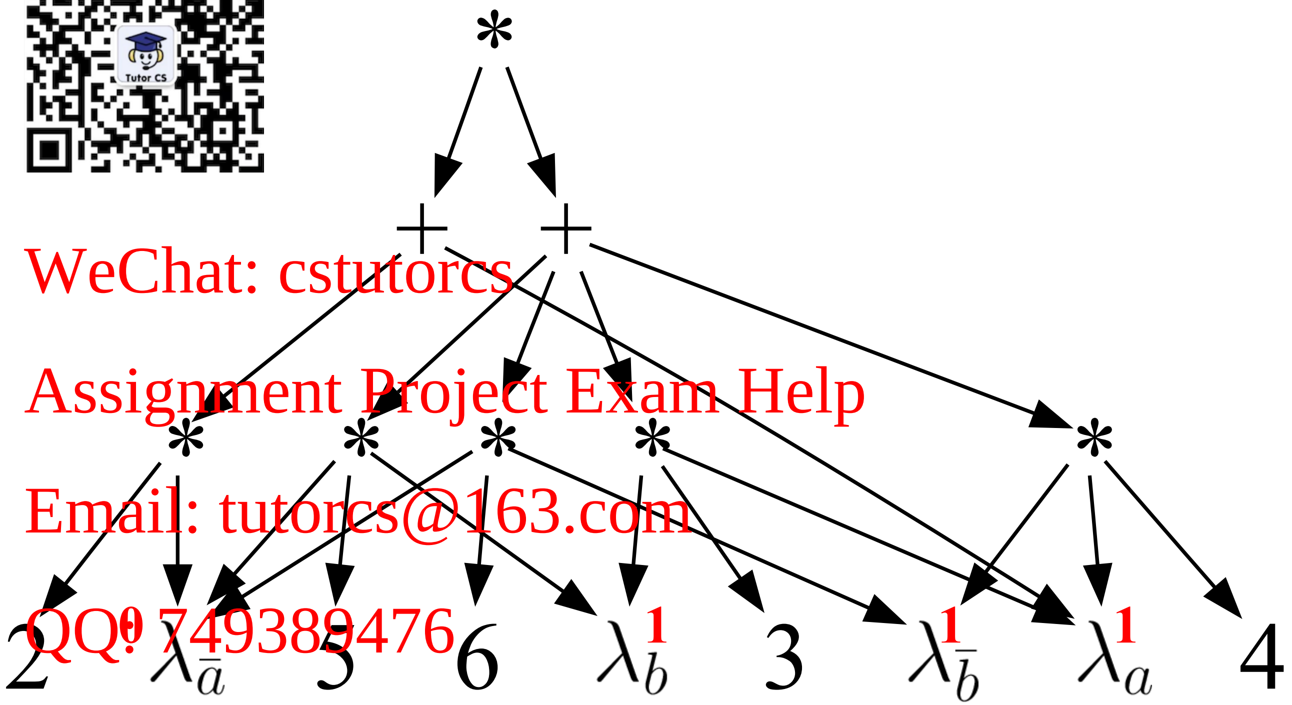


Evaluate at a

Does AC Compute Marginals?

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12

$$f(a) = 7$$



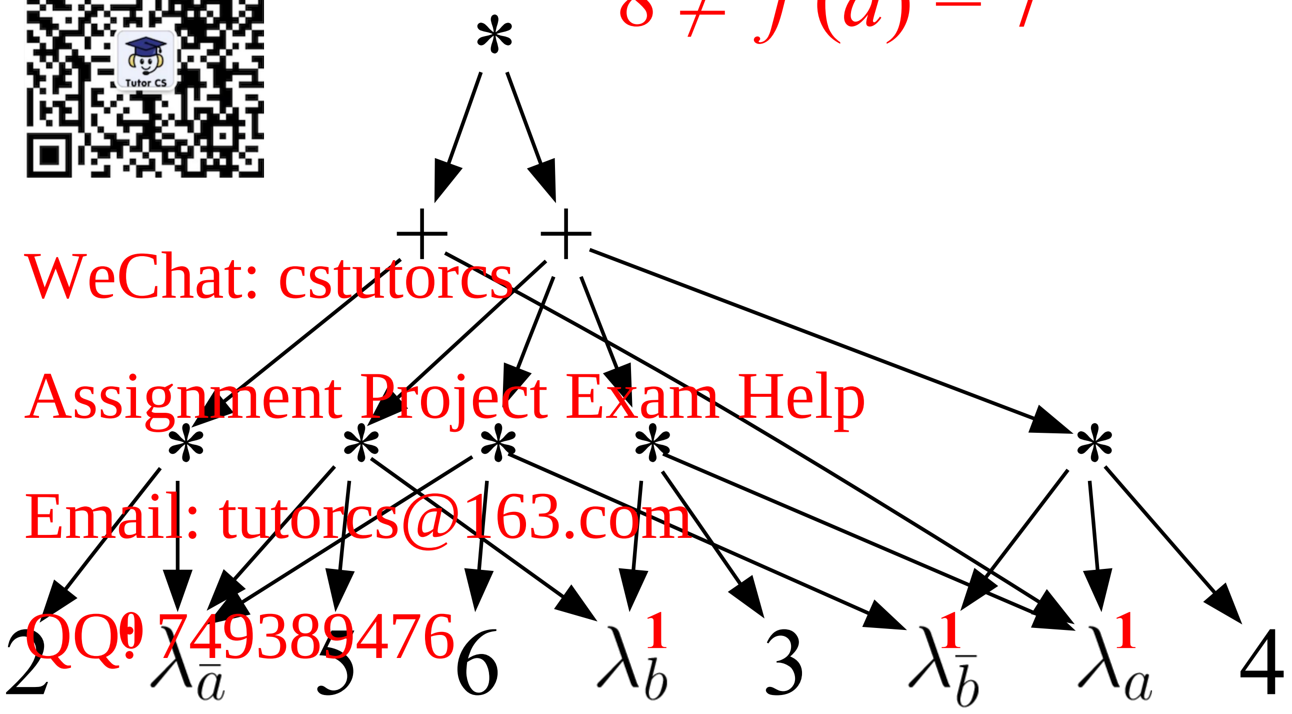
Evaluate at a

Does AC Compute Marginals?

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12

$f(a) = 7$

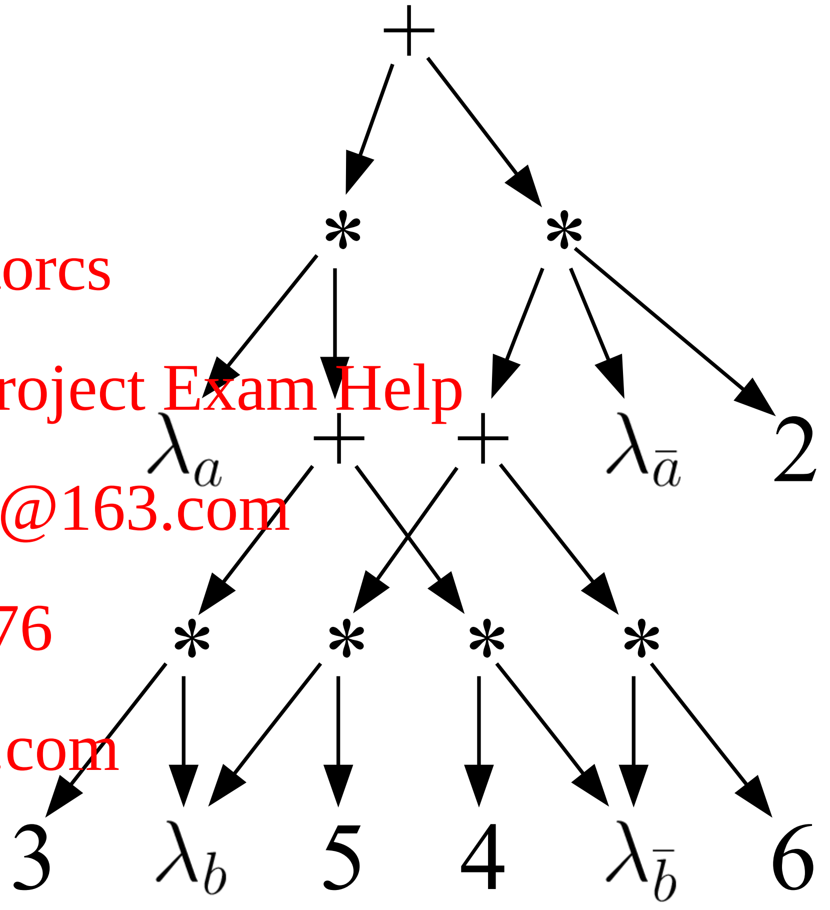
$8 \neq f(a) = 7$



Evaluate at a

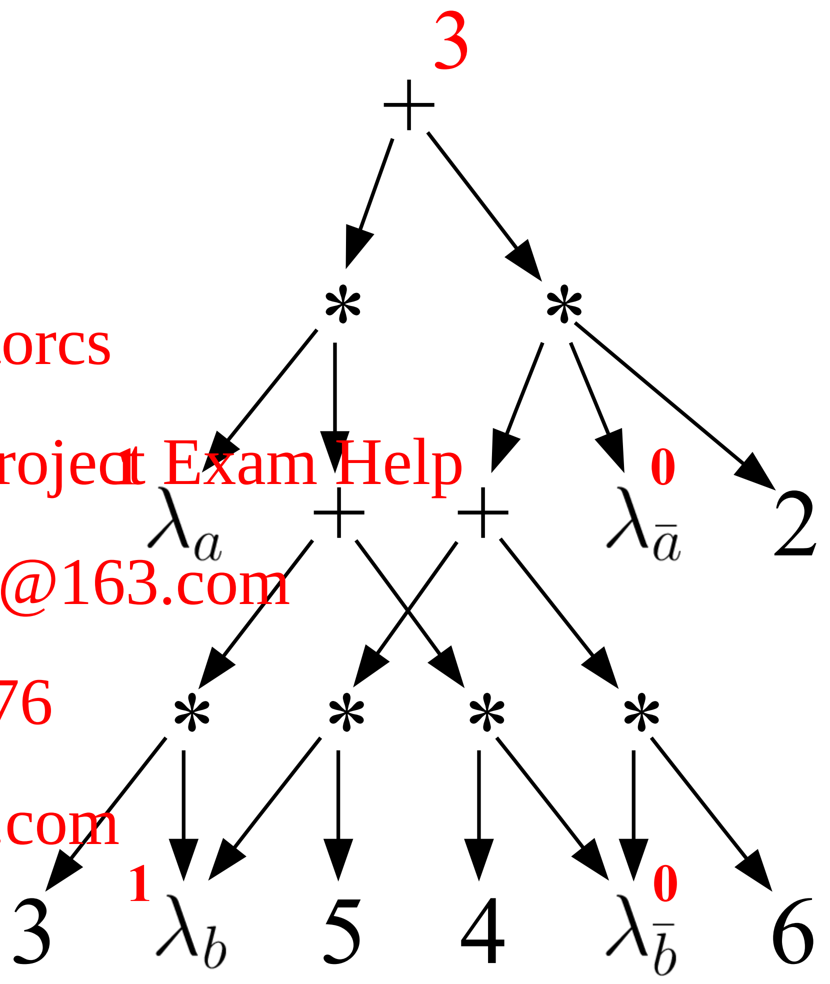
This AC Does Compute Marginals

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12



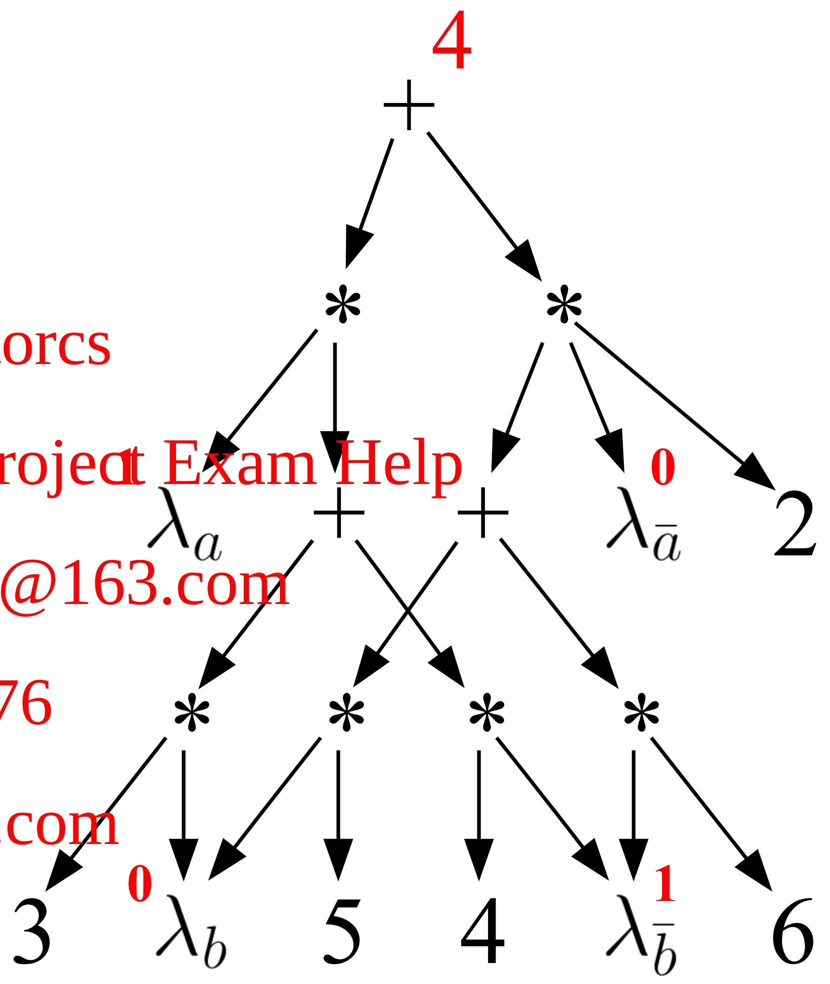
Factor Computed by AC

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12



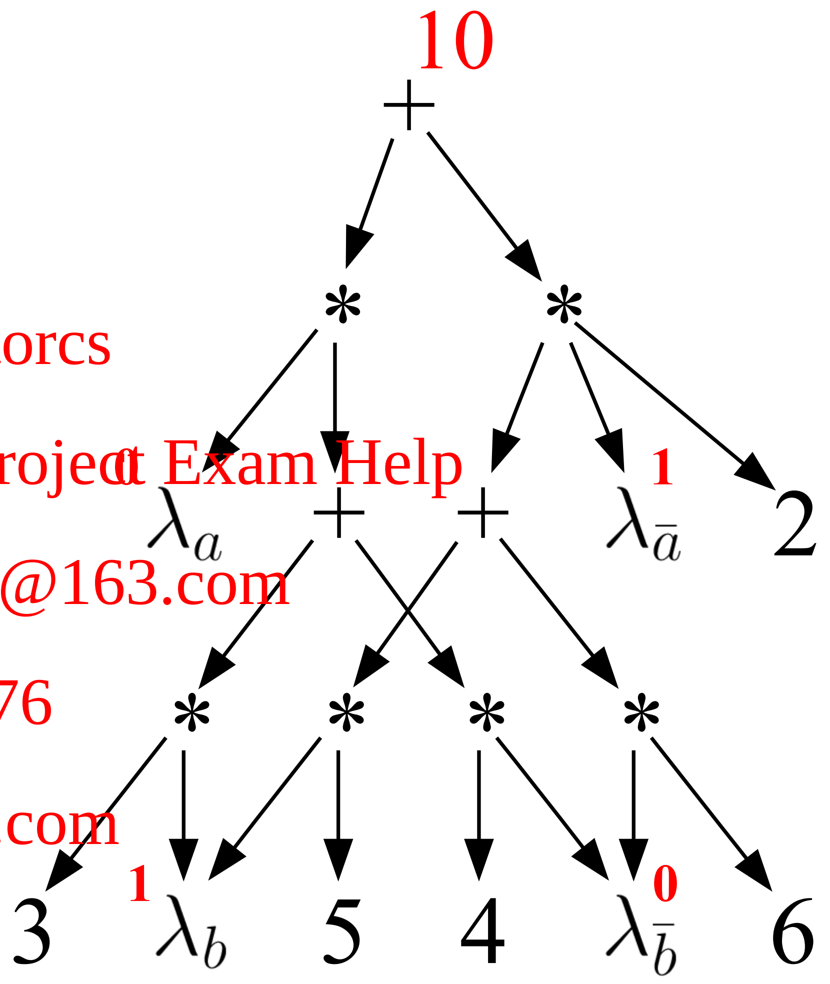
Factor Computed by AC

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12



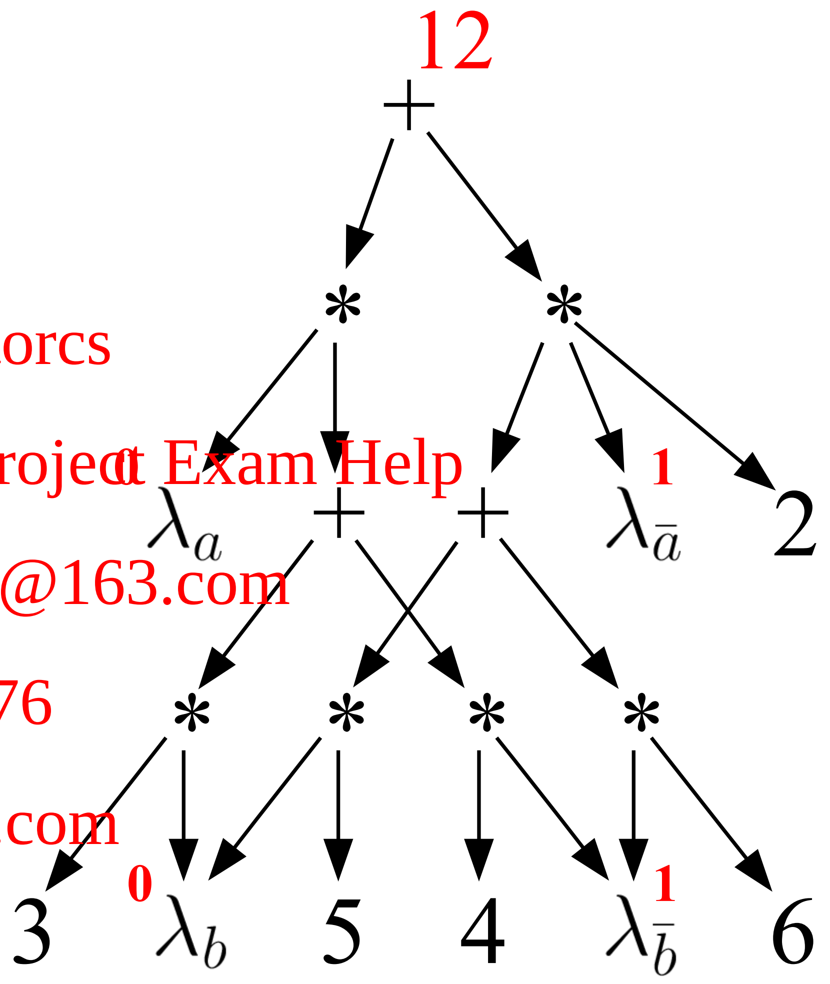
Factor Computed by AC

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12



Factor Computed by AC

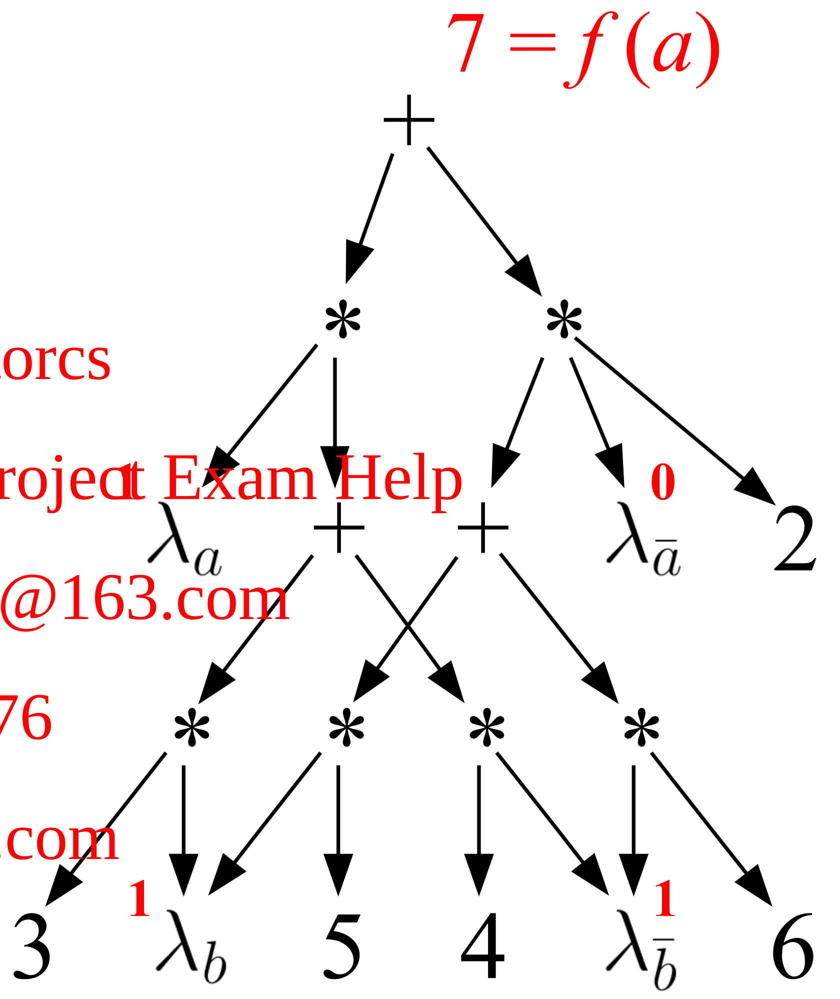
A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12



AC Computes Marginals

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12

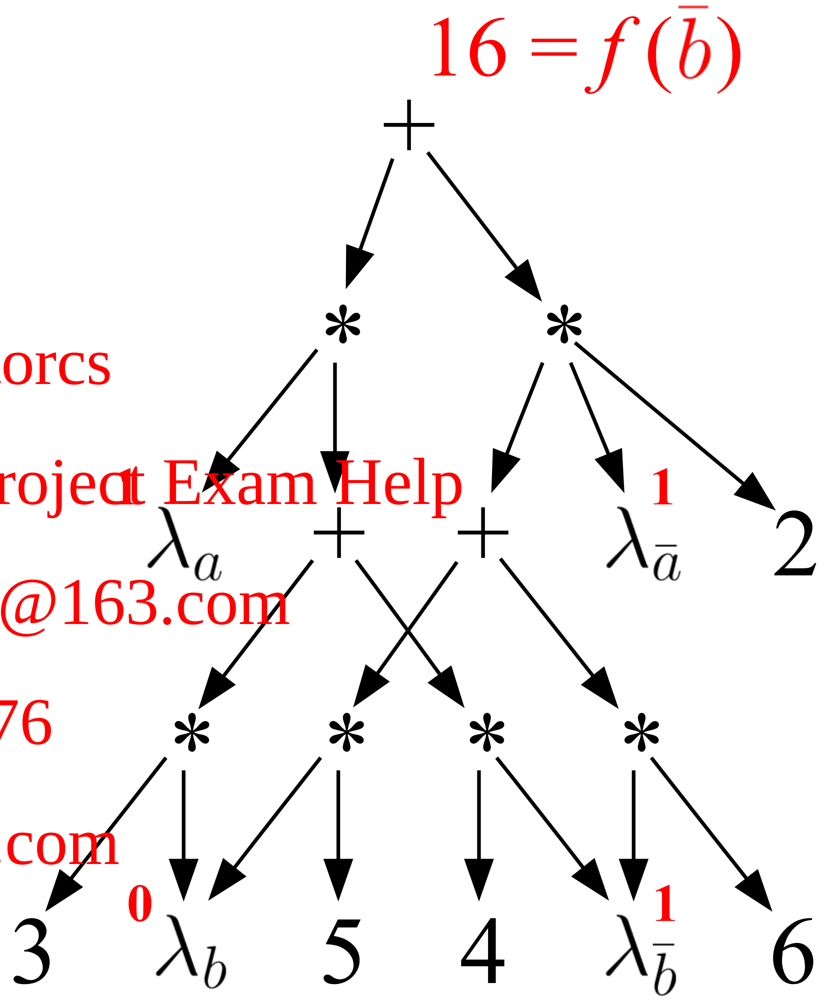
$$f(a) = 7$$



AC Computes Marginals

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12

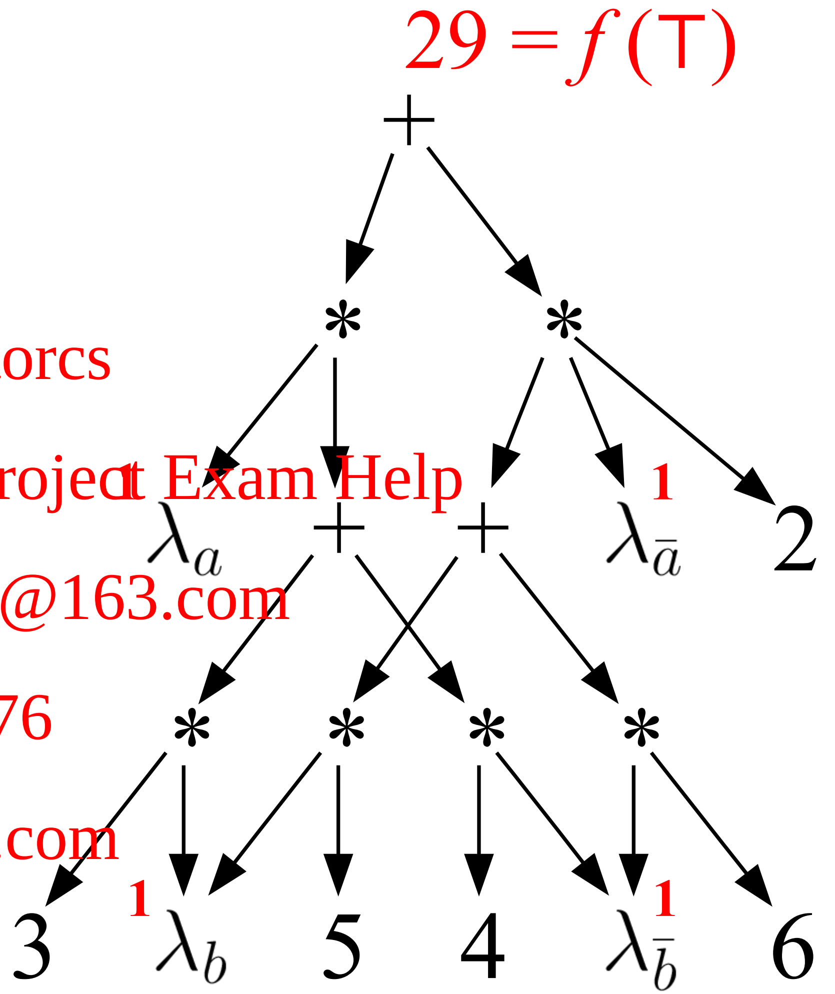
$$f(\bar{b}) = 16$$



AC Computes Marginals

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12

$$f(\top) = 29$$



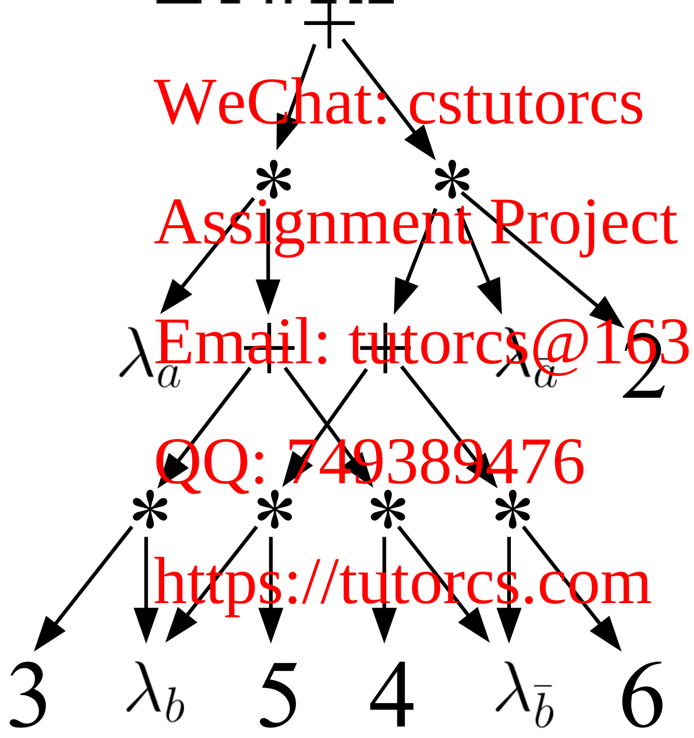
Why Did This AC Compute Marginals?

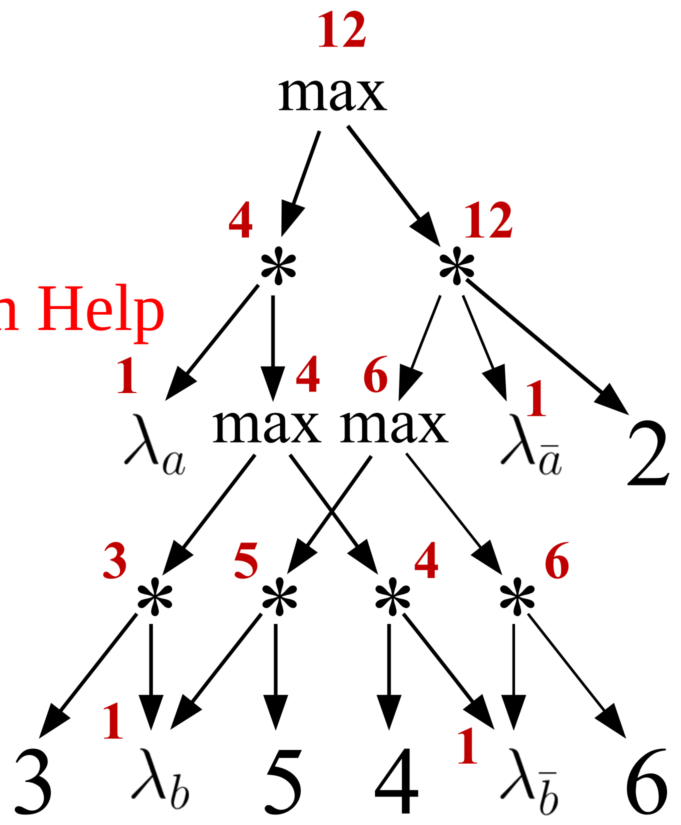
Three Fundamental Questions

- What factor f(**X**) does an AC **represent?**
  (the reference factor)
- When does an AC **compute marginals** of factor f(**X**)?
  (in linear time)
- When does an AC **compute MPE** of factor f(**X**)?
  (in linear time)

This AC Also Computes MPE

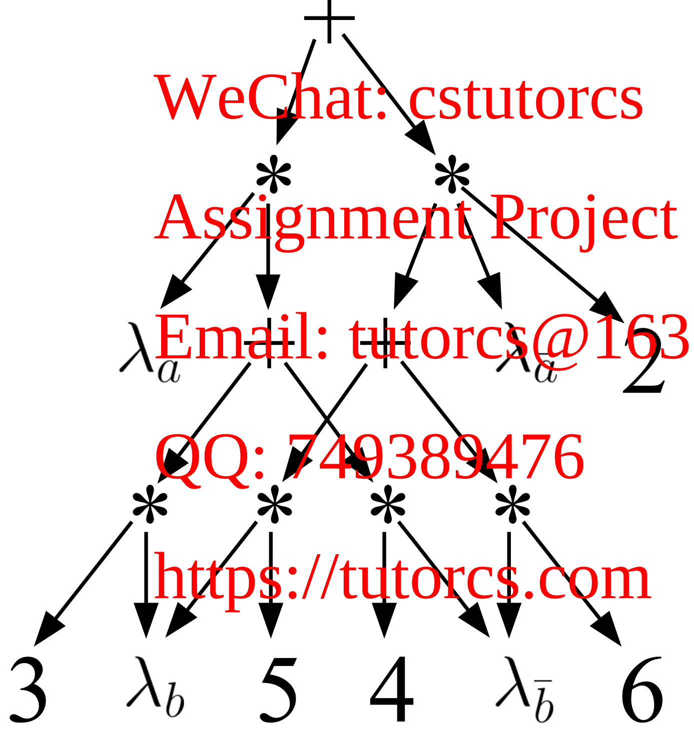
A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12

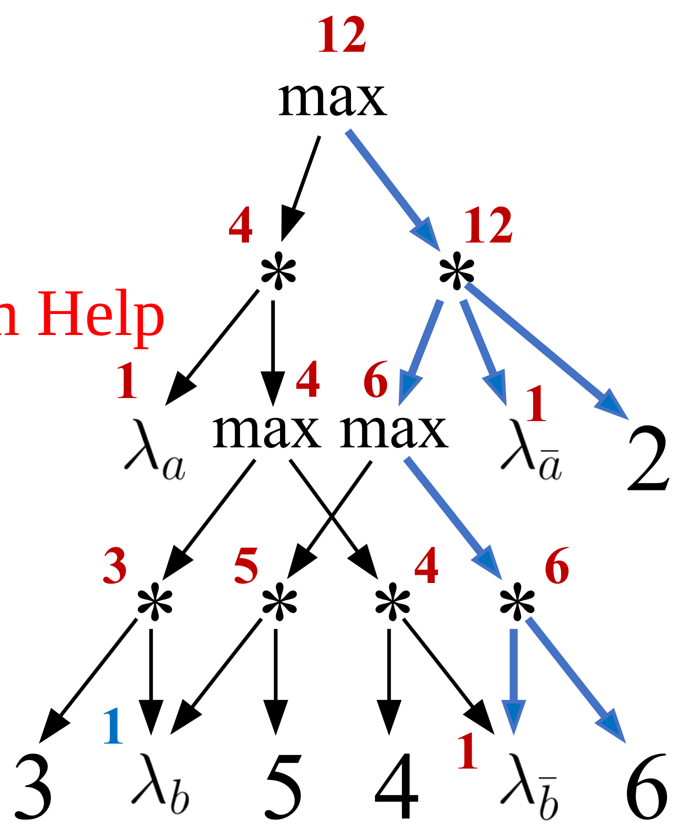




This AC Also Computes MPE

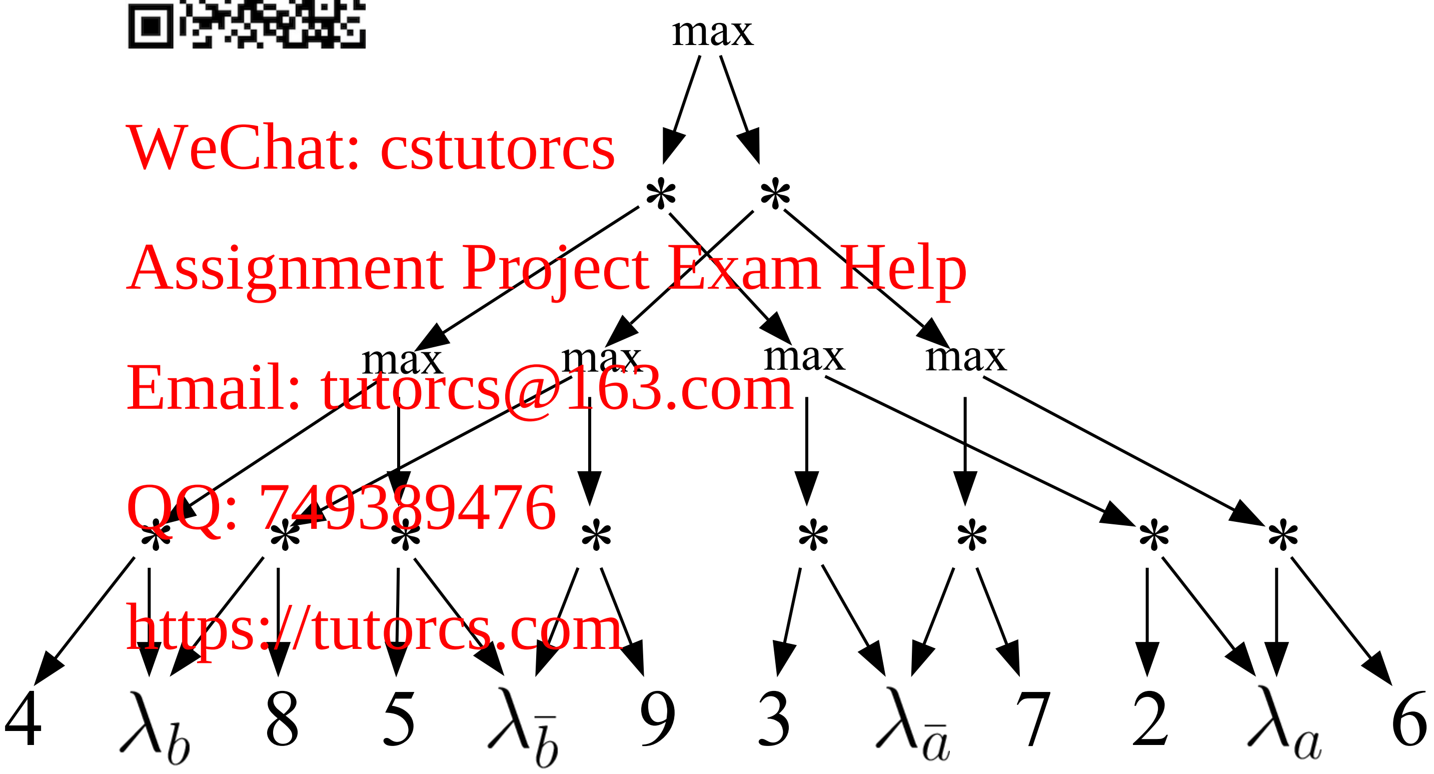
A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12





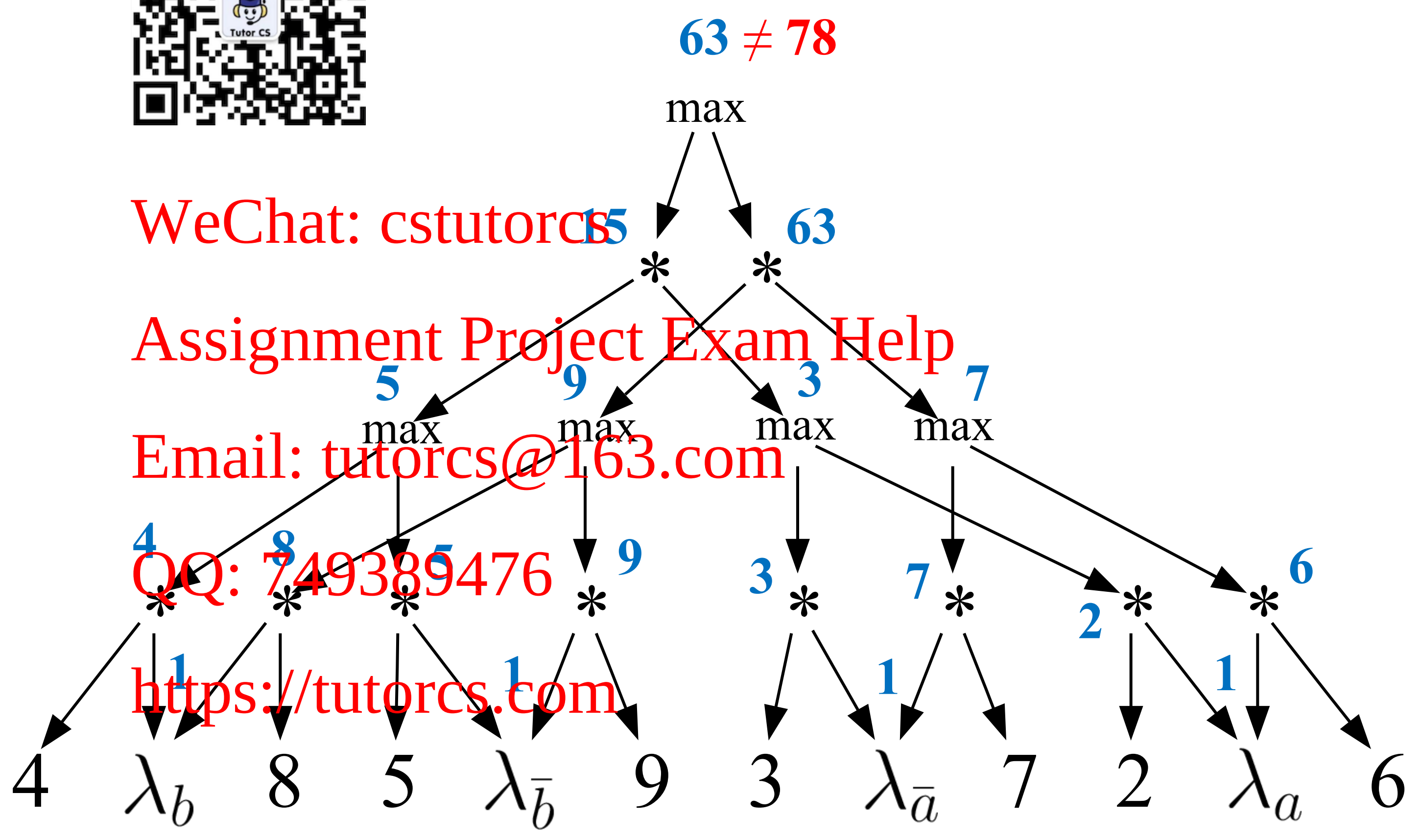
This AC Does Not Compute MPE

A	B	$f(A,B)$
a	b	56
a	$\bar{b}$	64
$\bar{a}$	b	68
$\bar{a}$	$\bar{b}$	78



This AC Does Not Compute MPE

A	B	$f(A,B)$
a	b	56
a	$\bar{b}$	64
$\bar{a}$	b	68
$\bar{a}$	$\bar{b}$	78

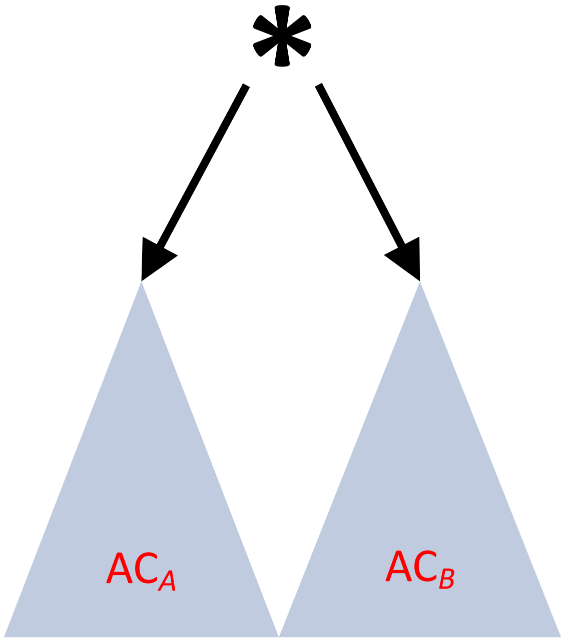


Why Did This AC Not Compute MPE?

Properties of Arithmetic Circuits

If an AC computes a factor *f*, then it will compute other quantities (marginals, MPE,…) depending on what properties it satisfies (decomposability, smoothness, determinism).

Properties of Arithmetic Circuits

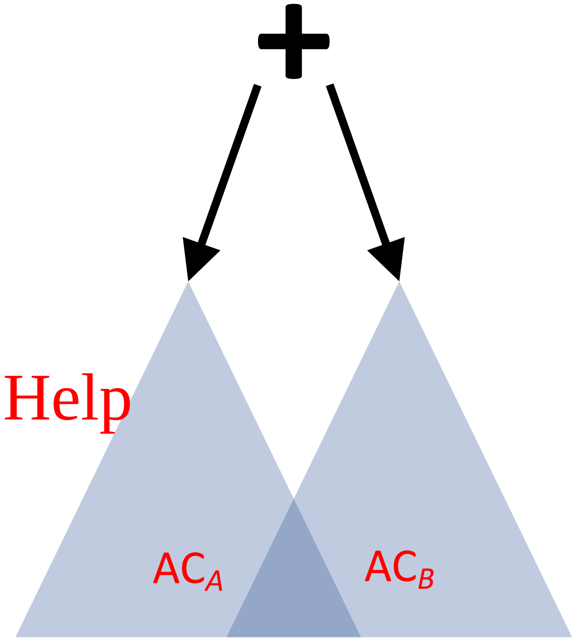


Decomposable

vars(AC_A) ∩ vars(AC_B) = ∅

Deterministic

at most one +-input is non-zero under any complete instantiation



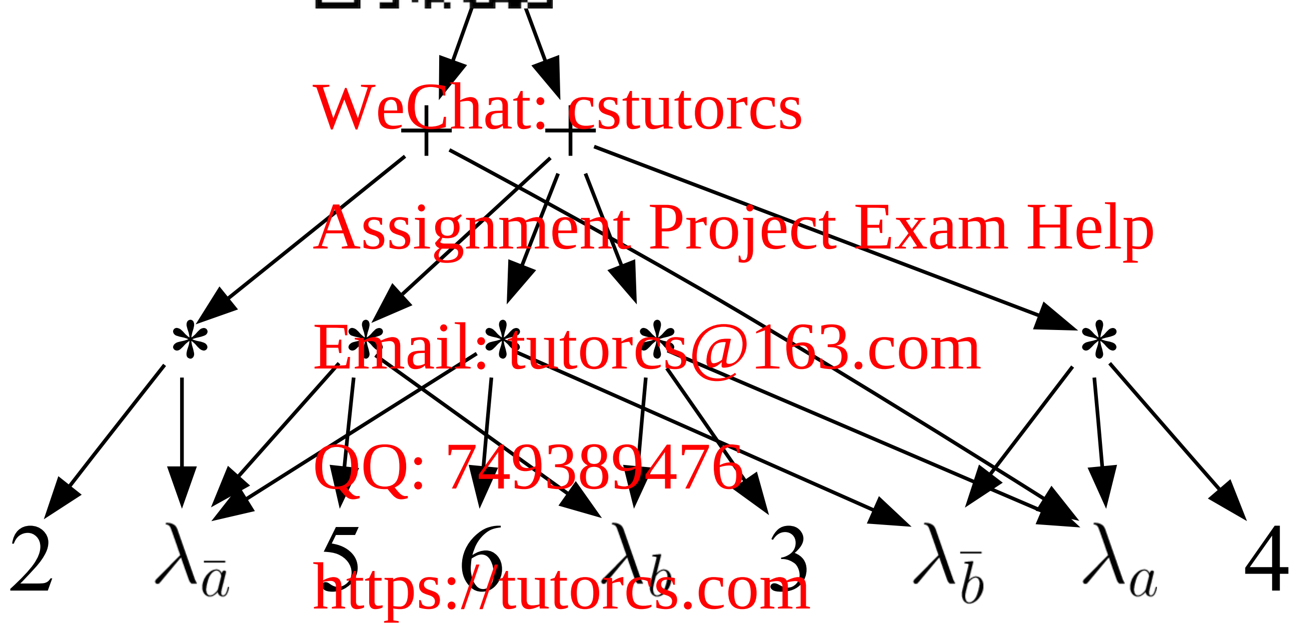
Smooth

vars(AC_A) = vars(AC_B)

Selective

Complete

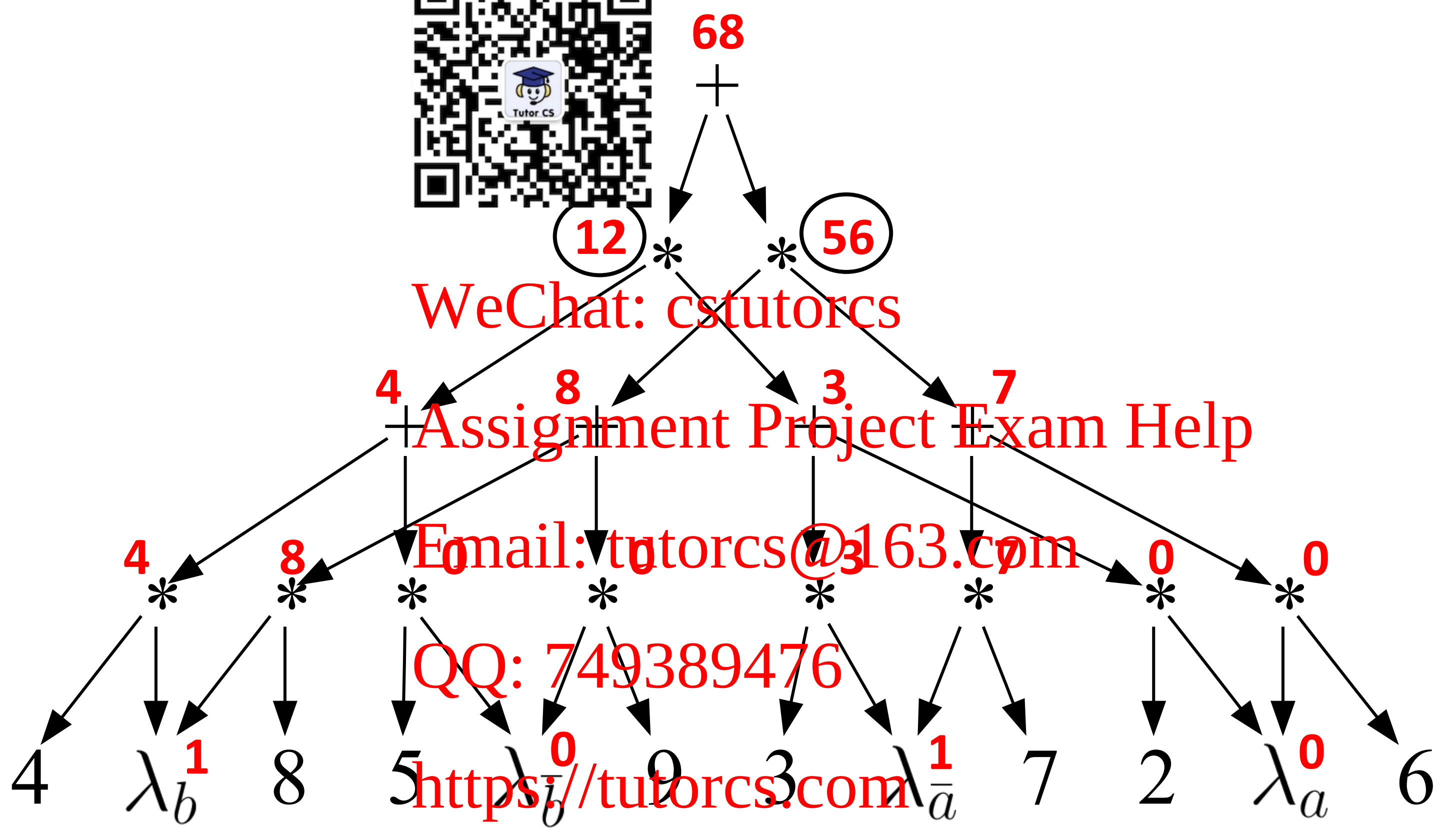
Not Decomposable



Not Decomposable

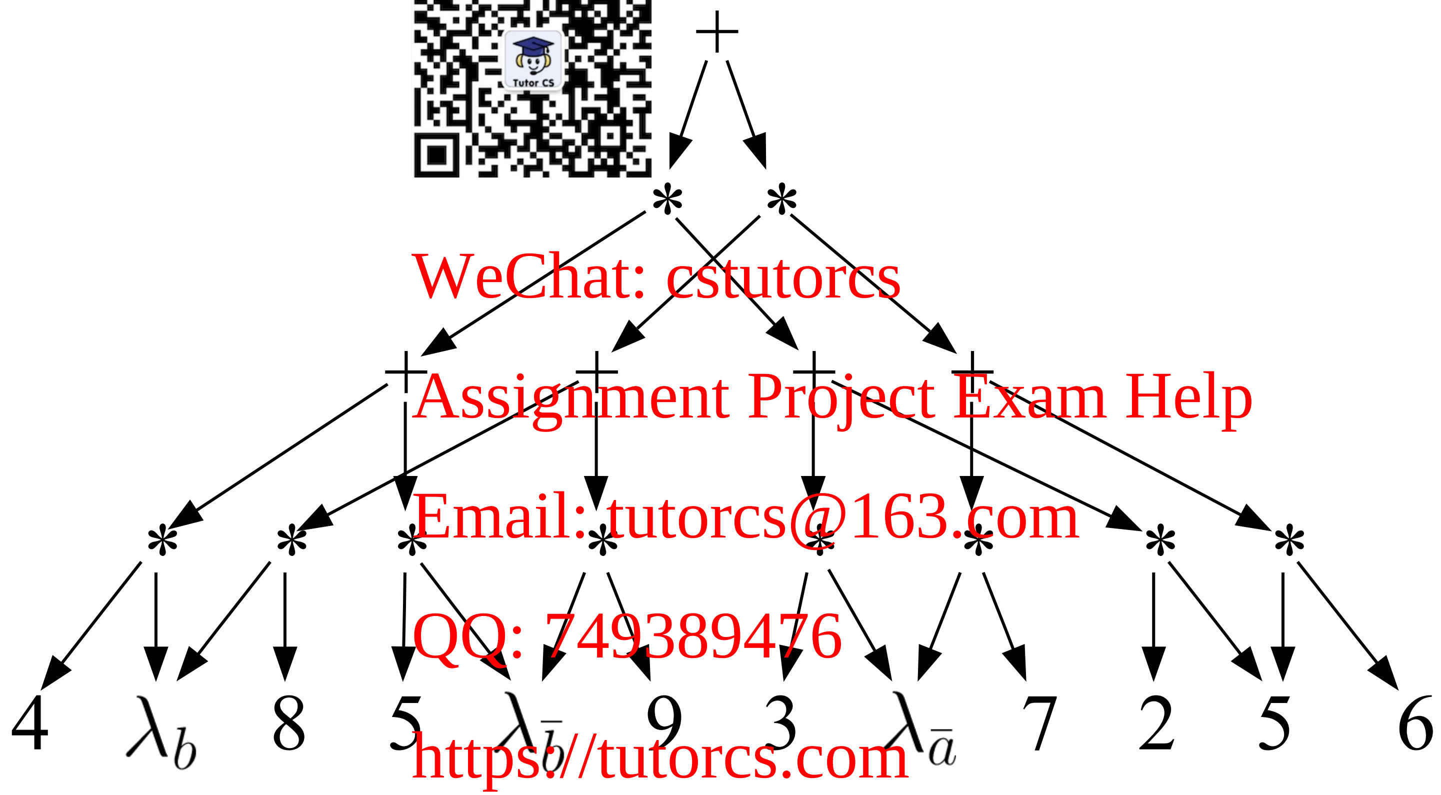
vars = {A}   **vars** = {A,B}

Not Deterministic

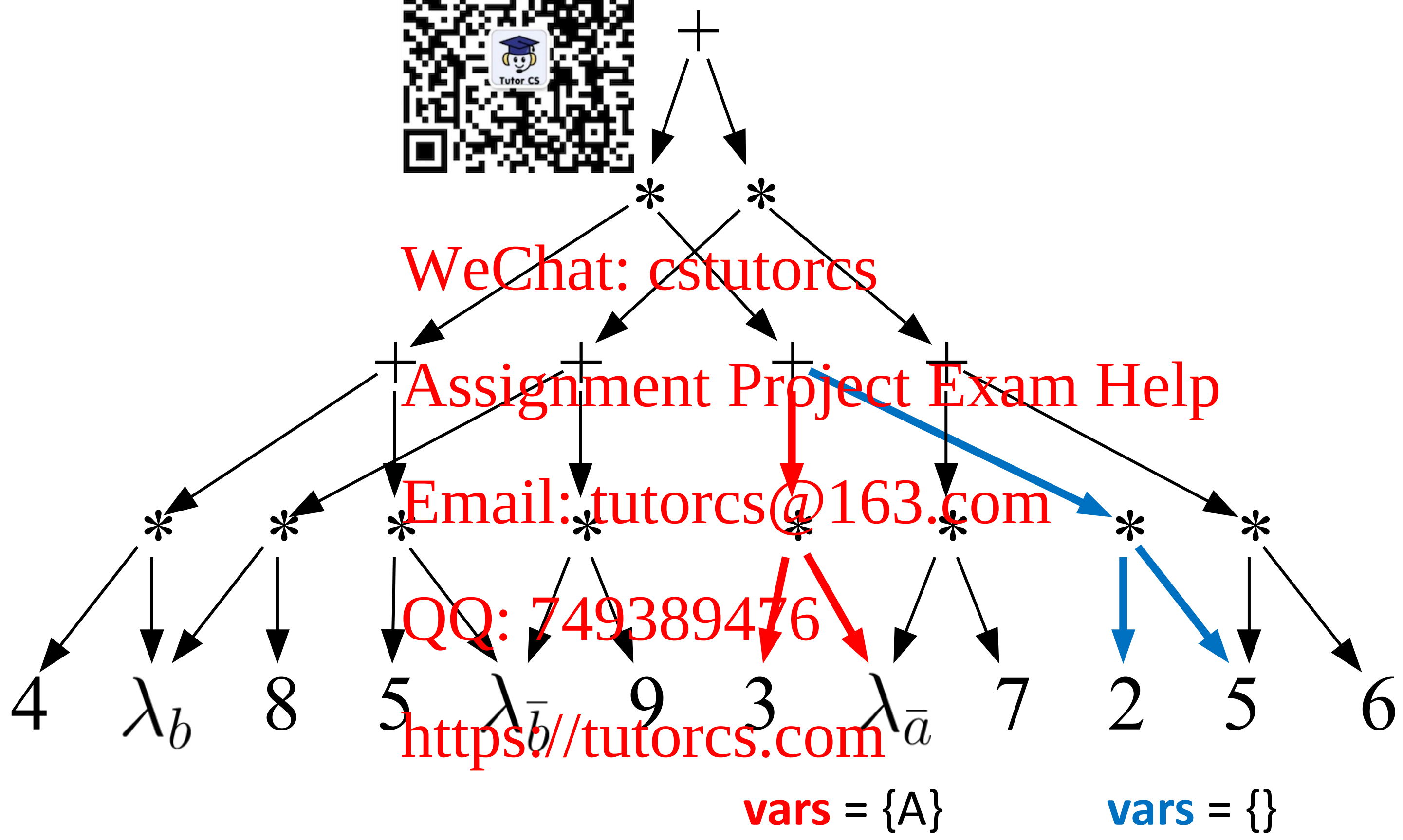


complete instantiation $\bar{a}b$

Not Smooth



Not Smooth



Computing Marginals and MPE

If an AC computes a factor f, then it will compute marginals and MPE (in linear time) if it is **decomposable**, **smooth** and **deterministic**.

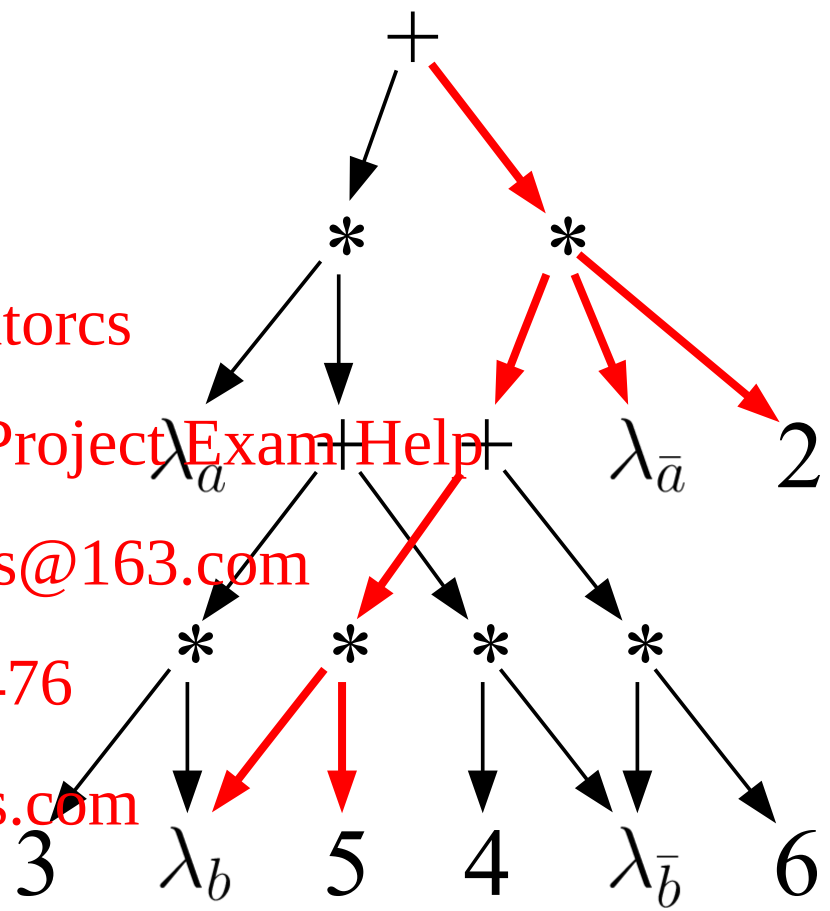
ACs: Marginals (2003), MPE (2006)

If an AC computes a factor f, then it will compute marginals if it is **decomposable** and **smooth**.

SPNs: Marginals (2011) **SPNs: MPE**

The Roles of Decomposability, Smoothness and Determinism

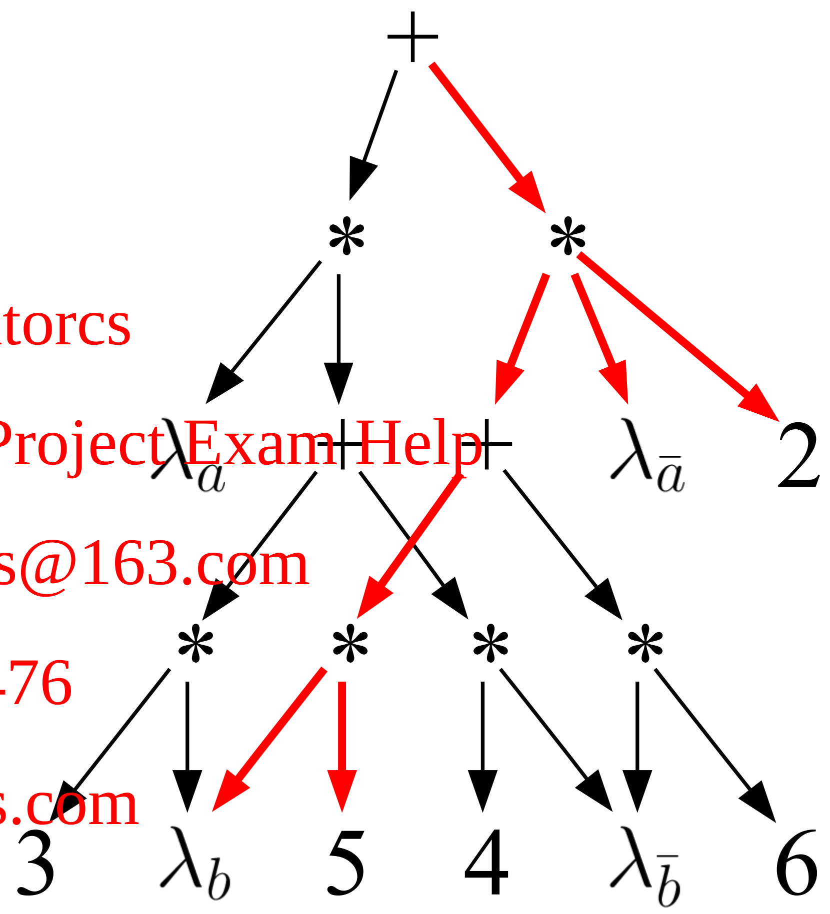
Fundamental Notion: Subcircuit



Fundamental Notion: Subcircuit

term: $\bar{a}b$

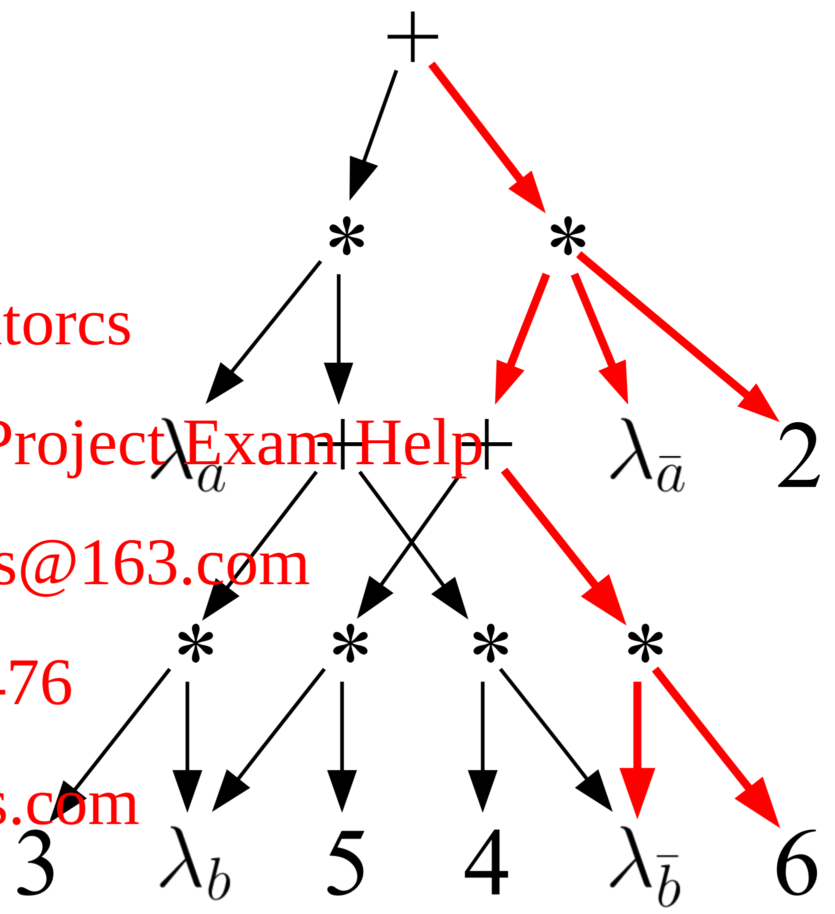
coefficient: $2*5 = 10$



Fundamental Notion: Subcircuit

term: $\bar{a}\,\bar{b}$

coefficient: $2*6 = 12$



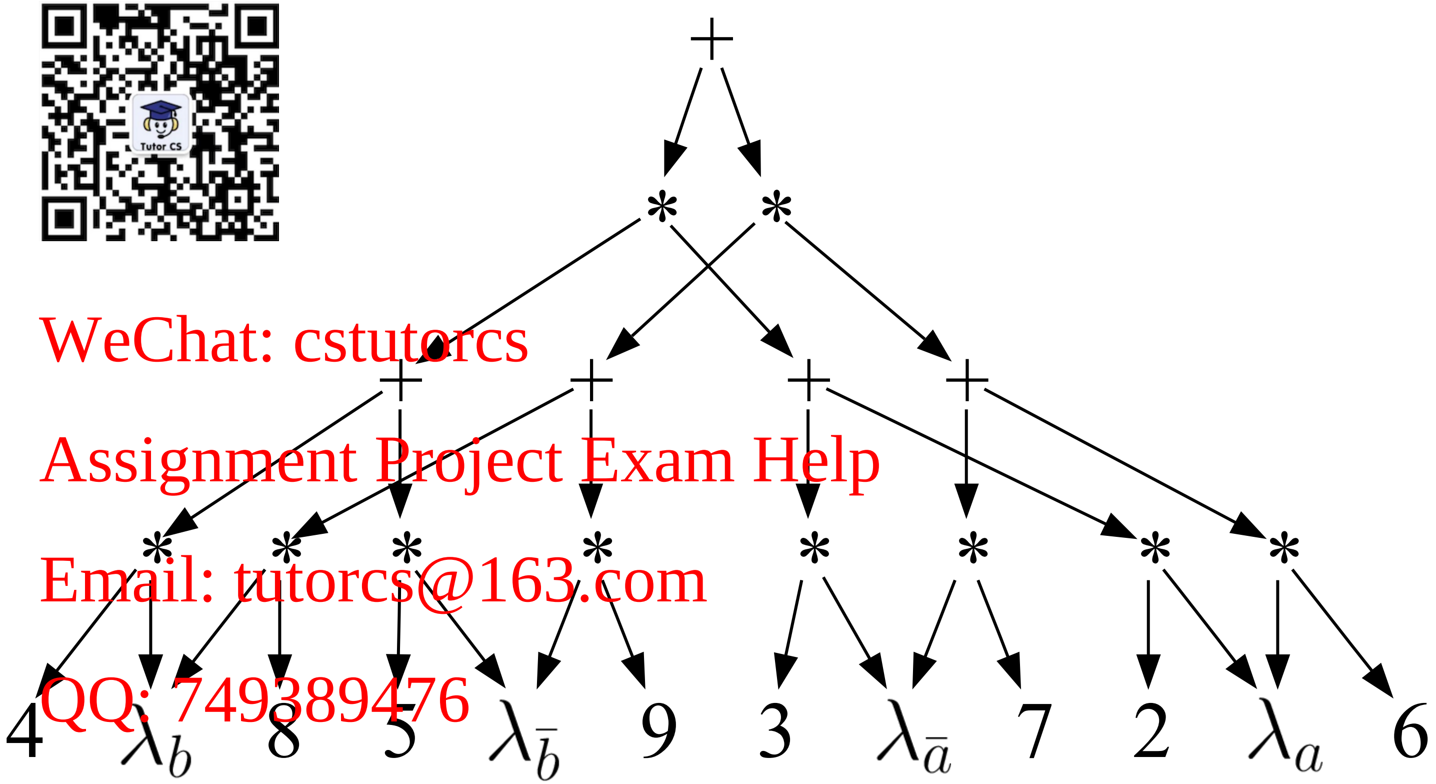
Decomposability and Smoothness

Decomposable →

subcircuit term is consistent

Smooth →

subcircuit term mentions all variables

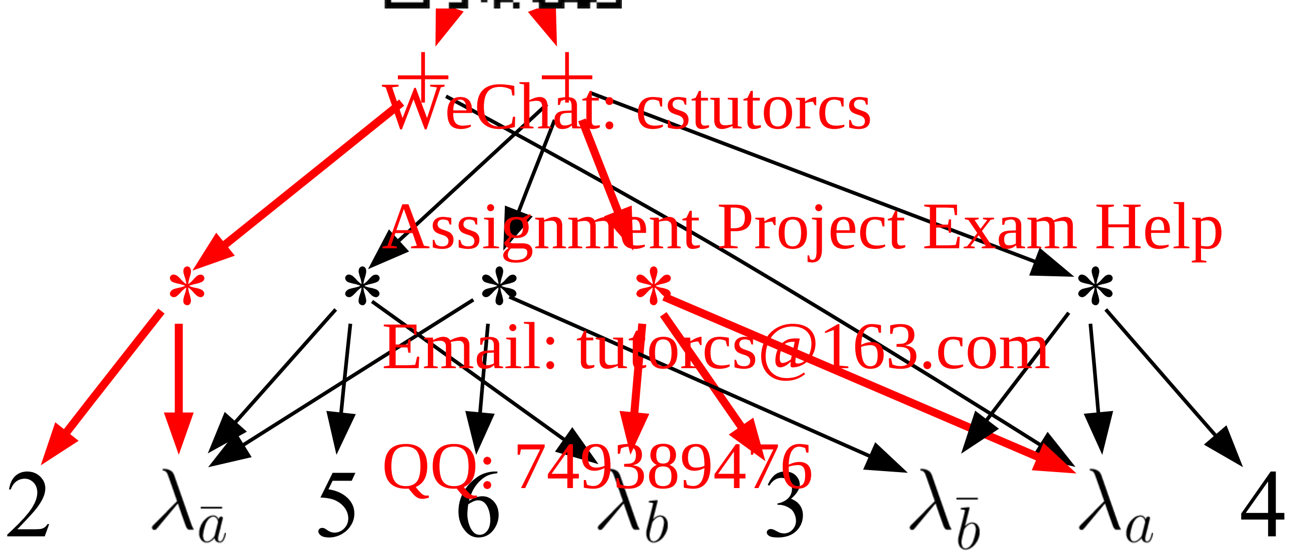


Decomposable and Smooth →

subcircuit term is a complete variable instantiation

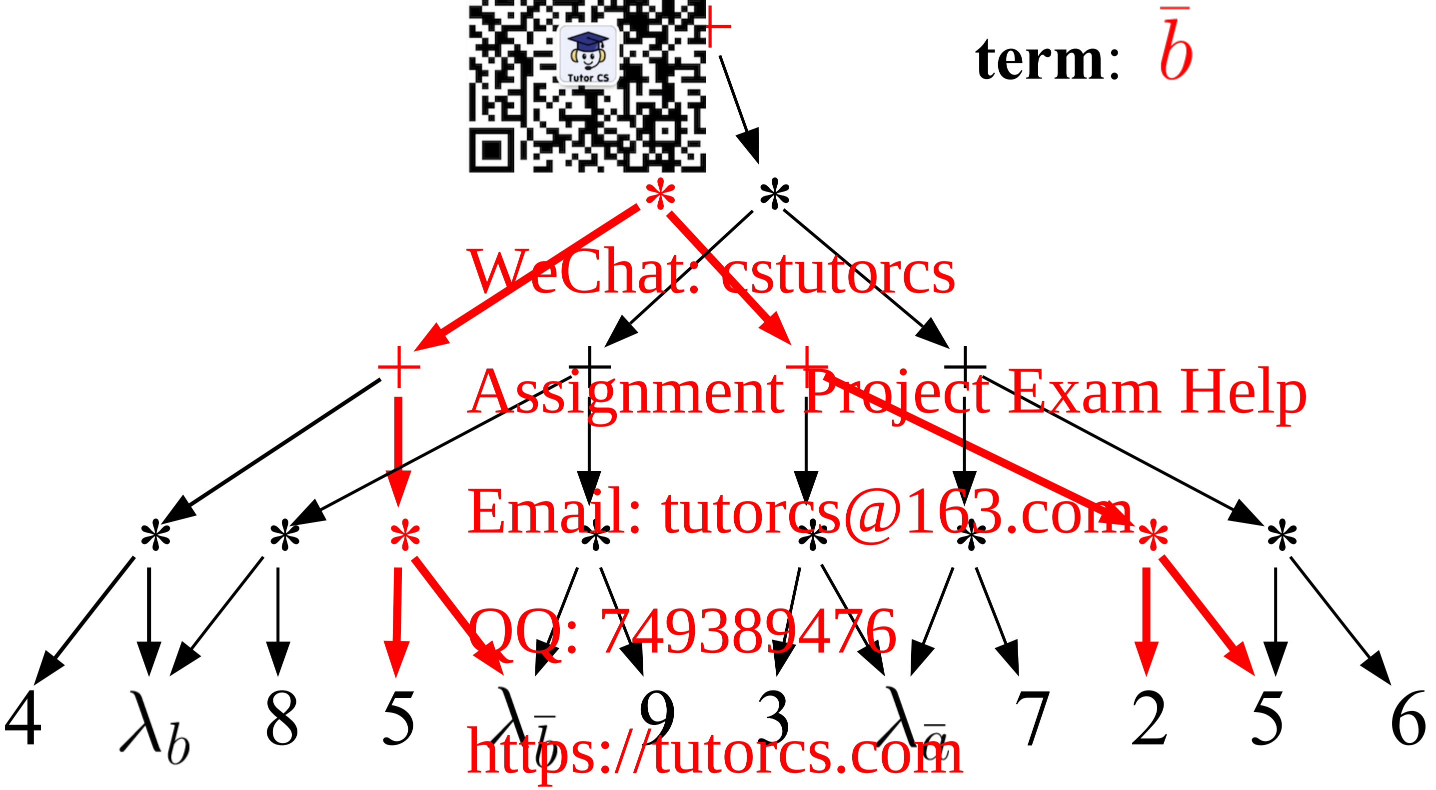
AC Not Decomposable

term: $\bar{a}\; b\; a$



subcircuit term is **not consistent**

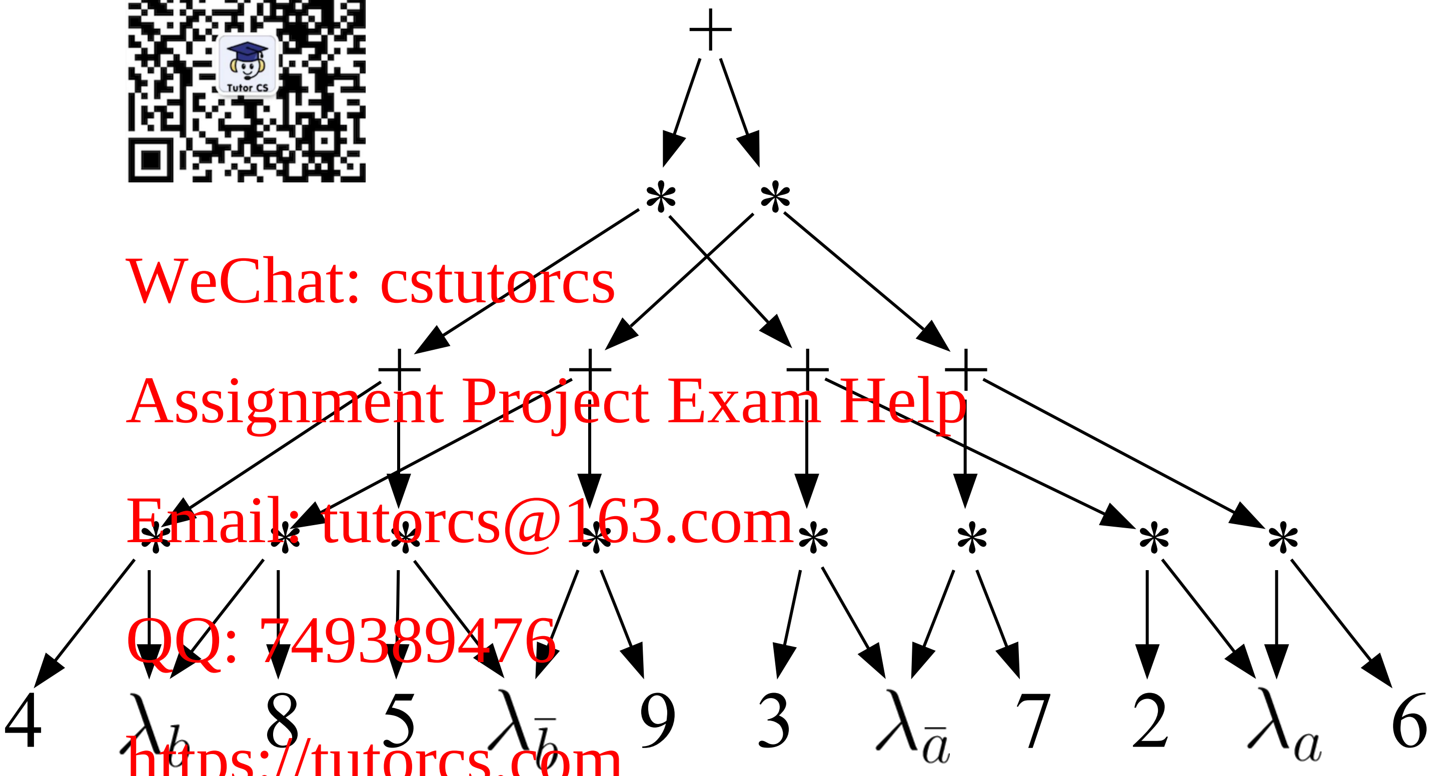
AC Not Smooth



subcircuit term **does not mention every variable**

Role of Decomposability and Smoothness

A	B	$f(A,B)$
a	b	56
a	$\bar{b}$	64
$\bar{a}$	b	68
$\bar{a}$	$\bar{b}$	78



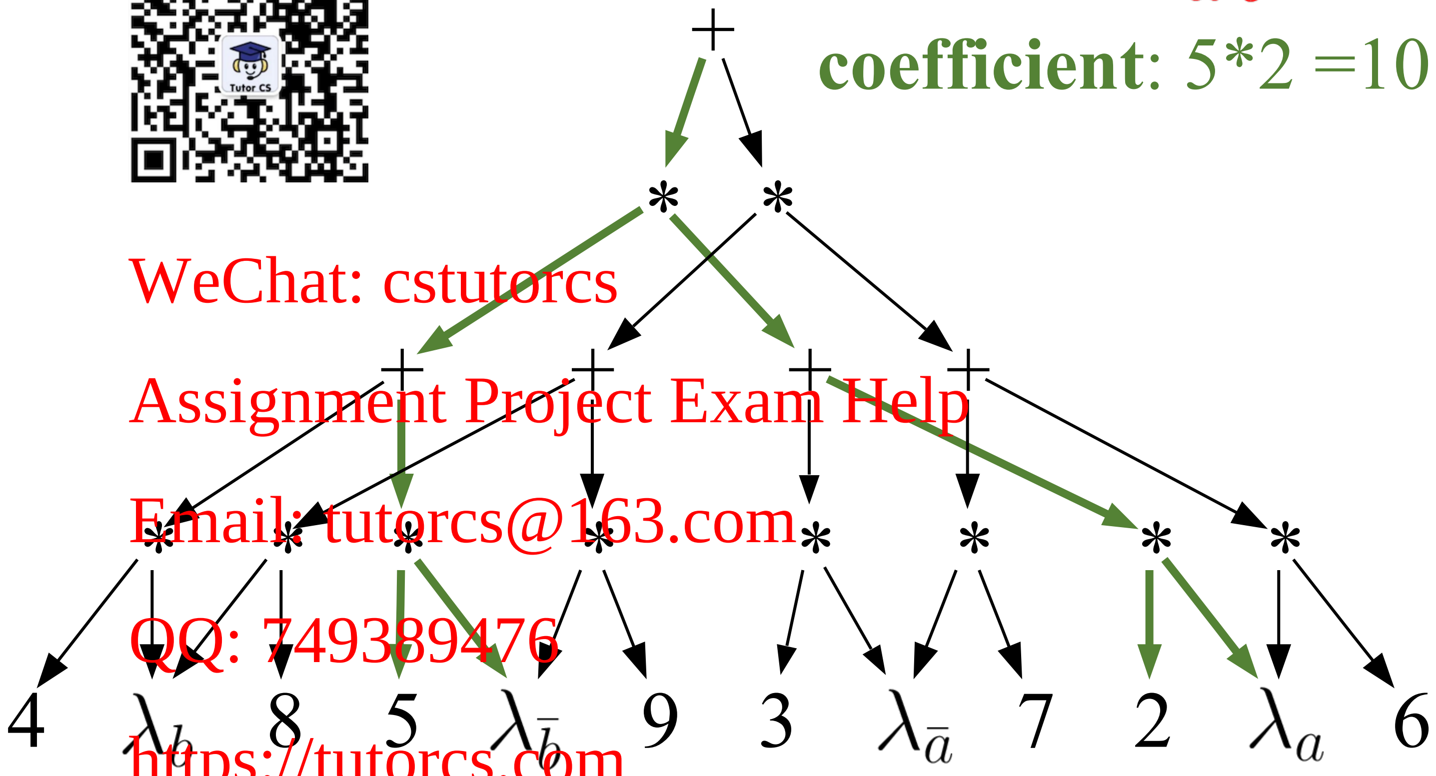
subcircuit term is a **complete variable instantiation**

Role of Decomposability and Smoothness

A	B	$f(A,B)$
a	b	56
a	$\bar{b}$	64
$\bar{a}$	b	68
$\bar{a}$	$\bar{b}$	78

term: $a\,\bar{b}$

coefficient: 5*2 =10



subcircuit term is a **complete variable instantiation**

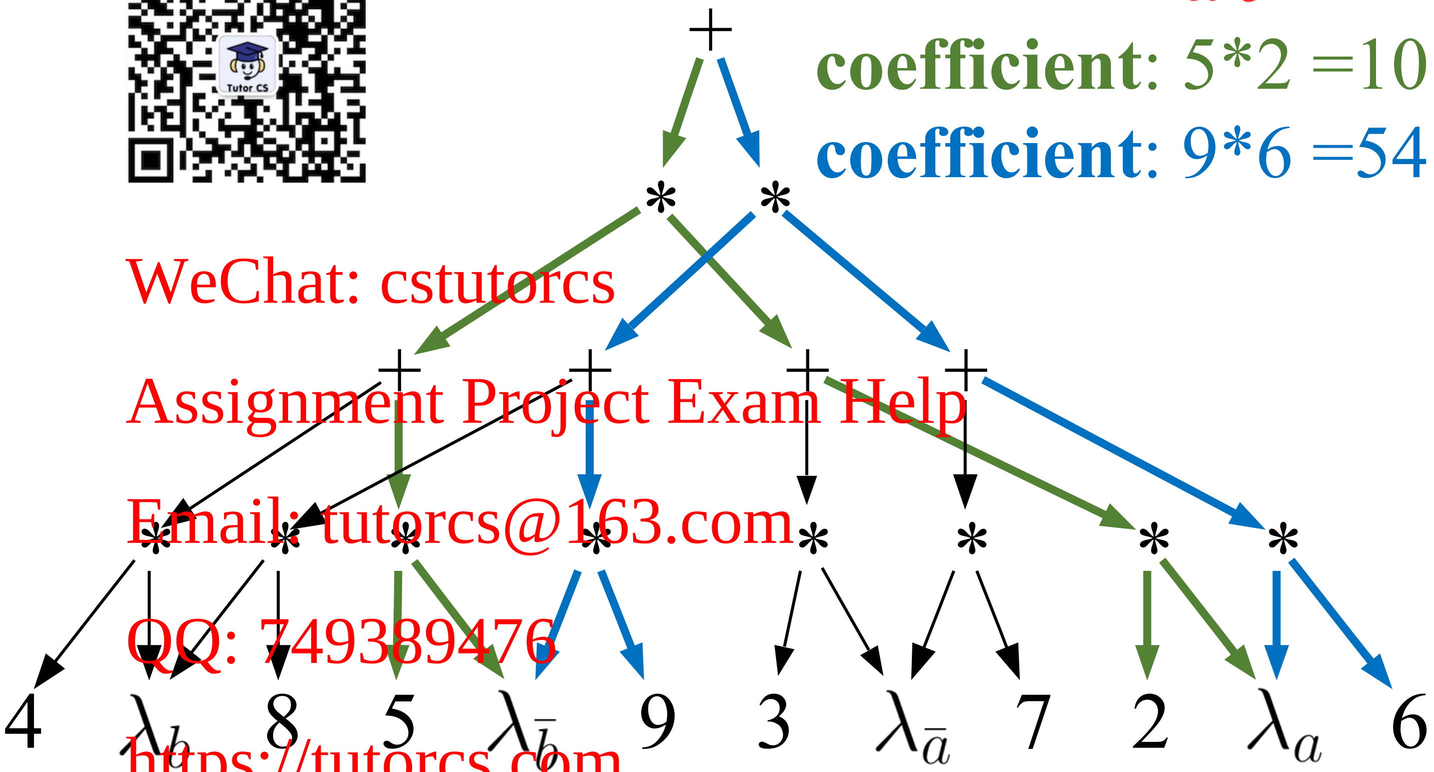
Role of Decomposability and Smoothness

A	B	$f(A,B)$
a	b	56
a	$\bar{b}$	64
$\bar{a}$	b	68
$\bar{a}$	$\bar{b}$	78

term: $a\,\bar{b}$

coefficient: 5*2 =10

coefficient: 9*6 =54

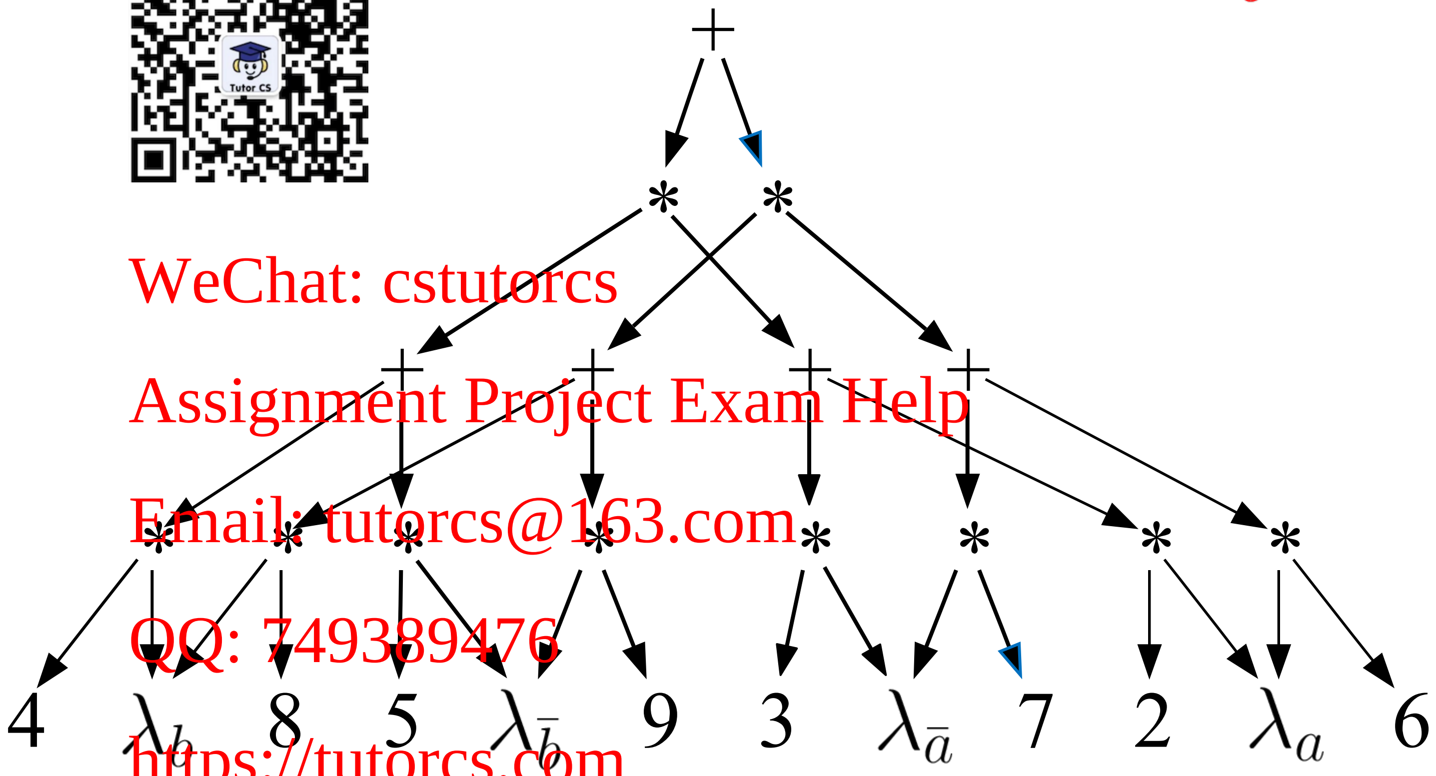


subcircuit term is a **complete variable instantiation**

Role of Decomposability and Smoothness

A	B	$f(A,B)$
a	b	56
a	$\bar{b}$	64
$\bar{a}$	b	68
$\bar{a}$	$\bar{b}$	78

term: $\bar{a}\bar{b}$



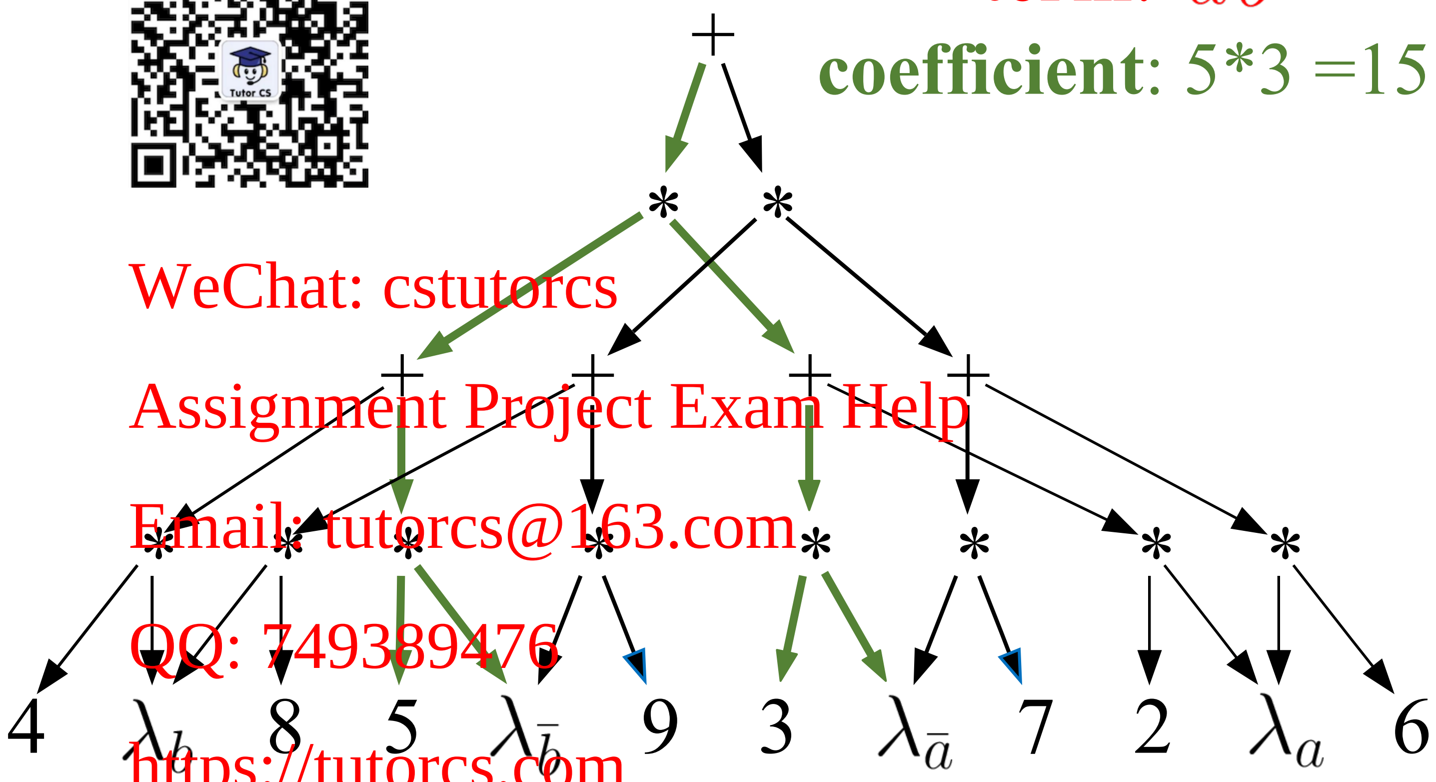
subcircuit term is a **complete variable instantiation**

Role of Decomposability and Smoothness

A	B	$f(A,B)$
a	b	56
a	$\bar{b}$	64
$\bar{a}$	b	68
$\bar{a}$	$\bar{b}$	78

term: $\bar{a}\bar{b}$

coefficient: 5*3 =15



subcircuit term is a **complete variable instantiation**

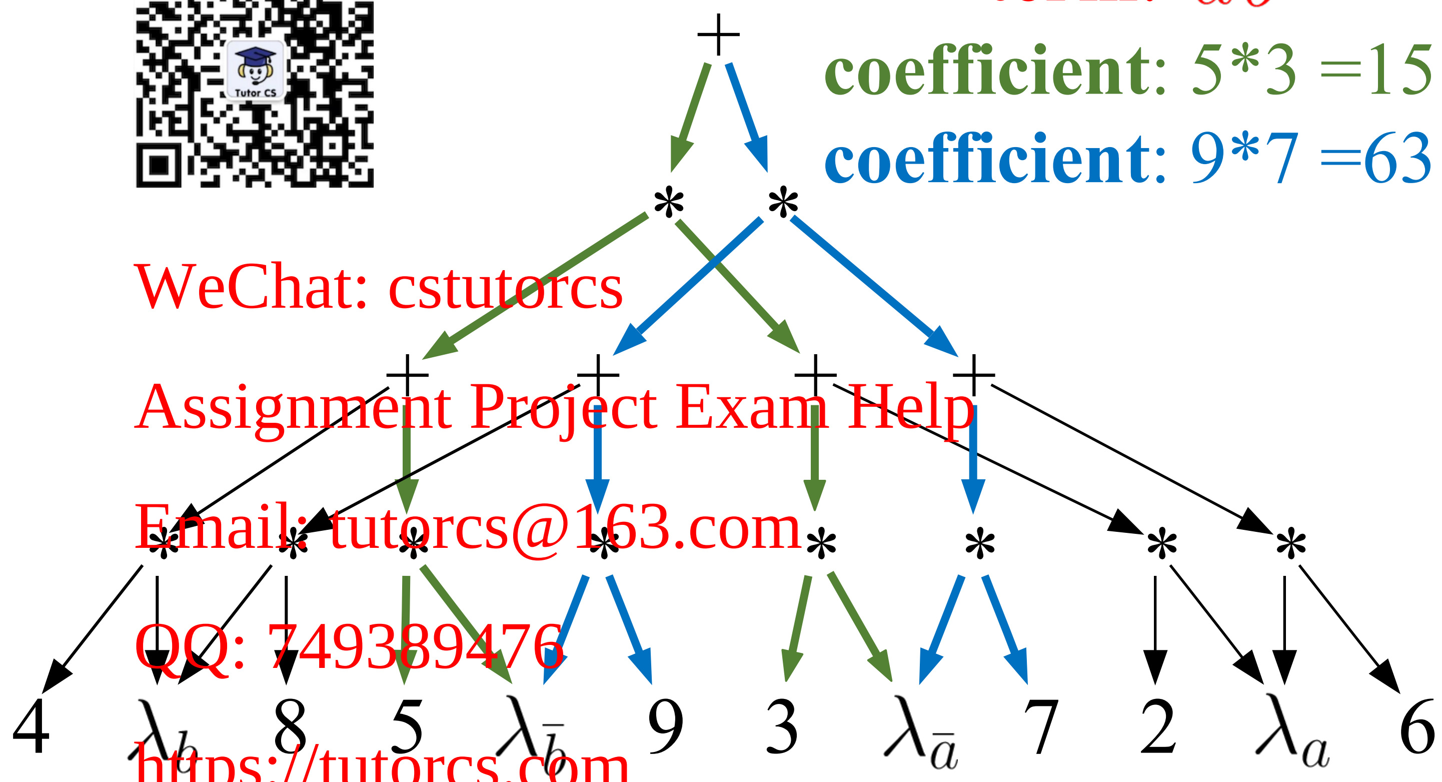
Role of Decomposability and Smoothness

A	B	$f(A,B)$
a	b	56
a	$\bar{b}$	64
$\bar{a}$	b	68
$\bar{a}$	$\bar{b}$	78

term: $\bar{a}\bar{b}$

coefficient: 5*3 =15

coefficient: 9*7 =63



subcircuit term is a **complete variable instantiation**

Role of Determinism

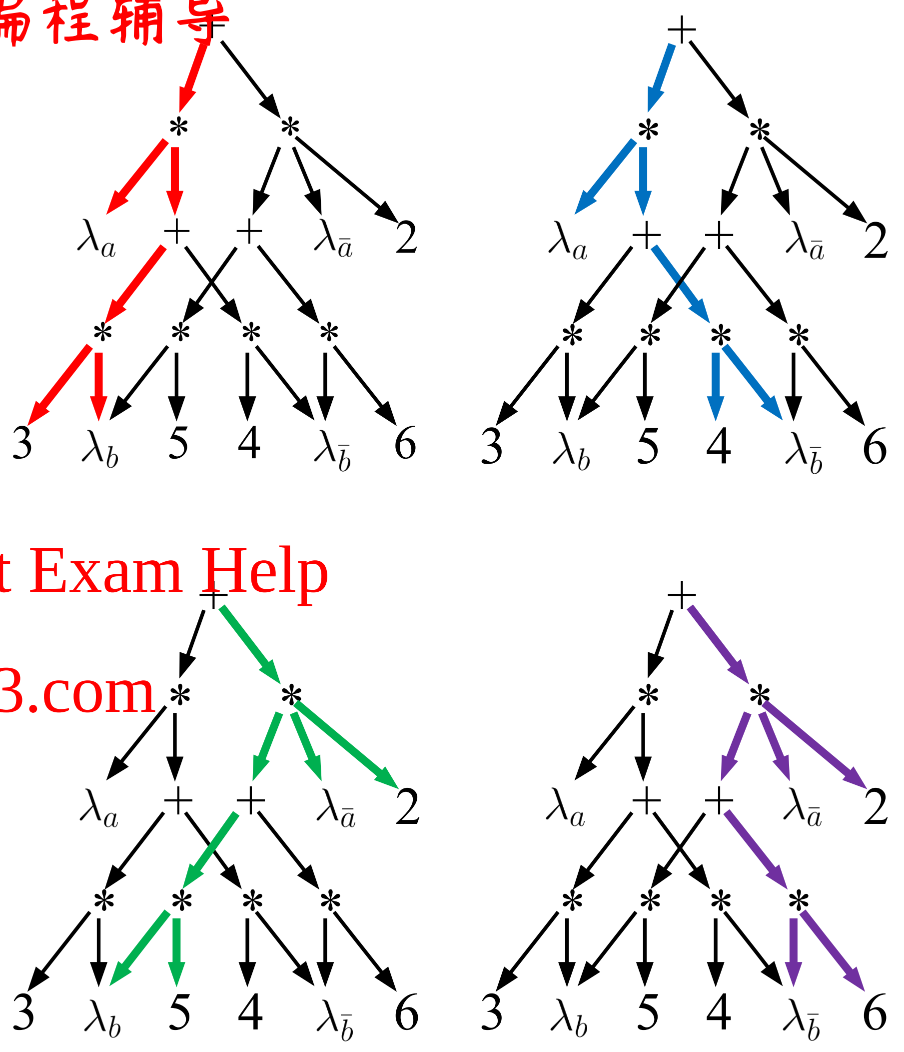
Decomposable & Smooth

subcircuit term is a complete variable instantiation

Determinism →

1-to-1 mapping between subcircuits and complete variable instantiations

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12



excluding subcircuits with zero coefficients

Understanding the MAR & MPE Algorithms

Setting indicator values to 0/1 modifies coefficients of subcircuits: **zeros out coefficients** of subcircuits that are not compatible with evidence

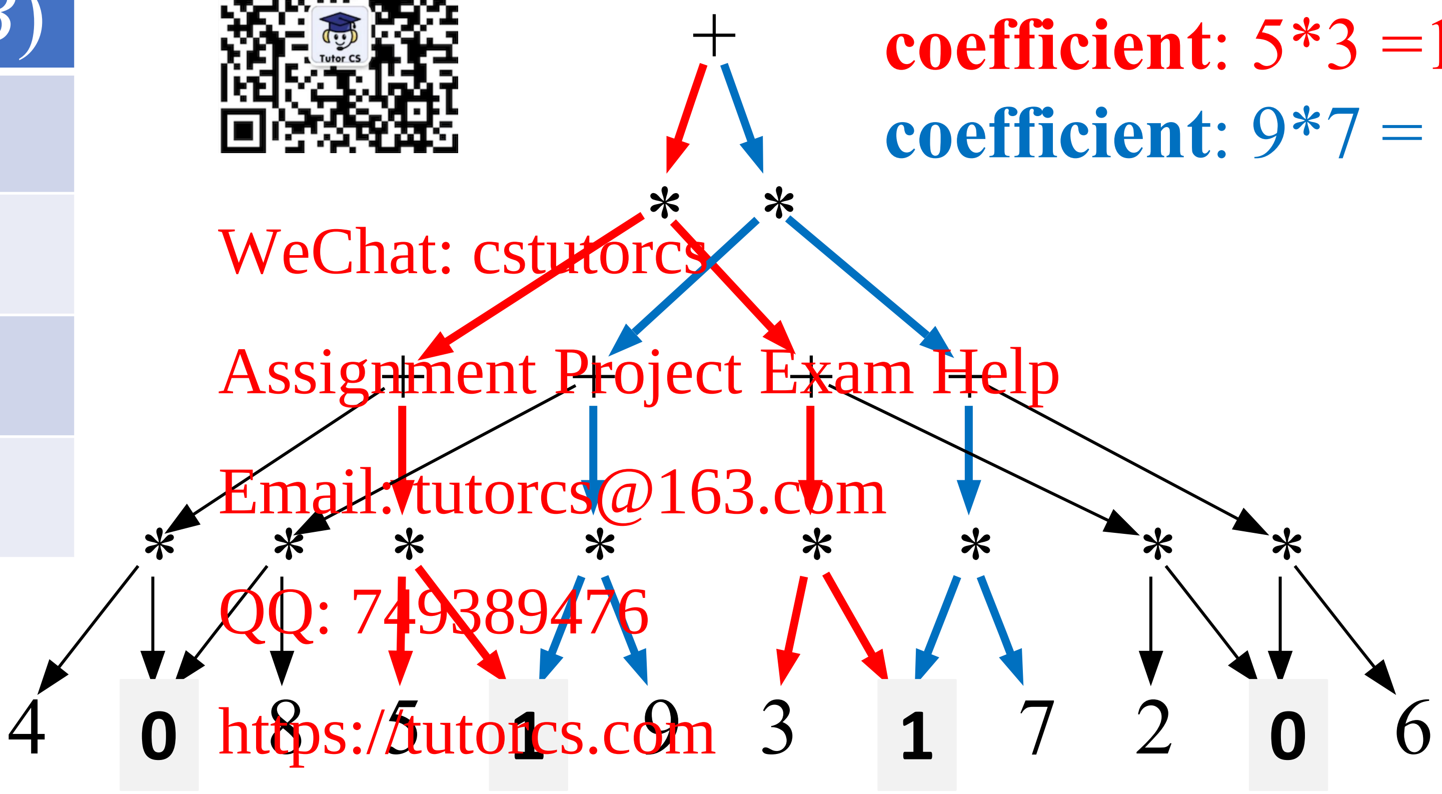
Marginals Algorithm

A	B	$f(A,B)$
a	b	56
a	$\bar{b}$	64
$\bar{a}$	b	68
$\bar{a}$	$\bar{b}$	78

term: $\bar{a}\,\bar{b}$

coefficient: 5*3 =15

coefficient: 9*7 = 63



sums coefficients of subcircuits that are compatible with evidence

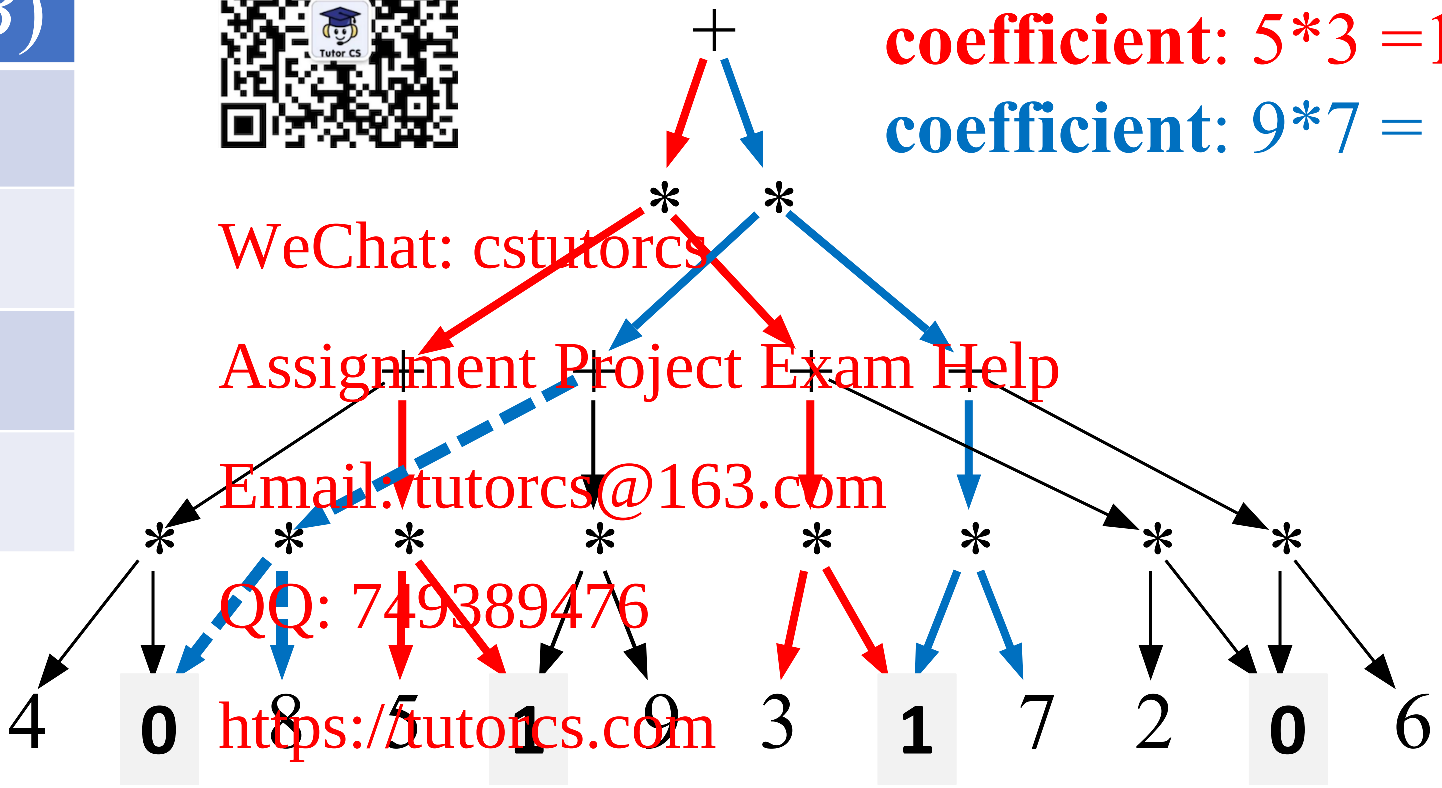
Marginals Algorithm

A	B	$f(A,B)$
a	b	56
a	$\bar{b}$	64
$\bar{a}$	b	68
$\bar{a}$	$\bar{b}$	78

term: $\bar{a}\,\bar{b}$

coefficient: 5*3 =15

coefficient: 9*7 = 63



sums coefficients of subcircuits that are compatible with evidence

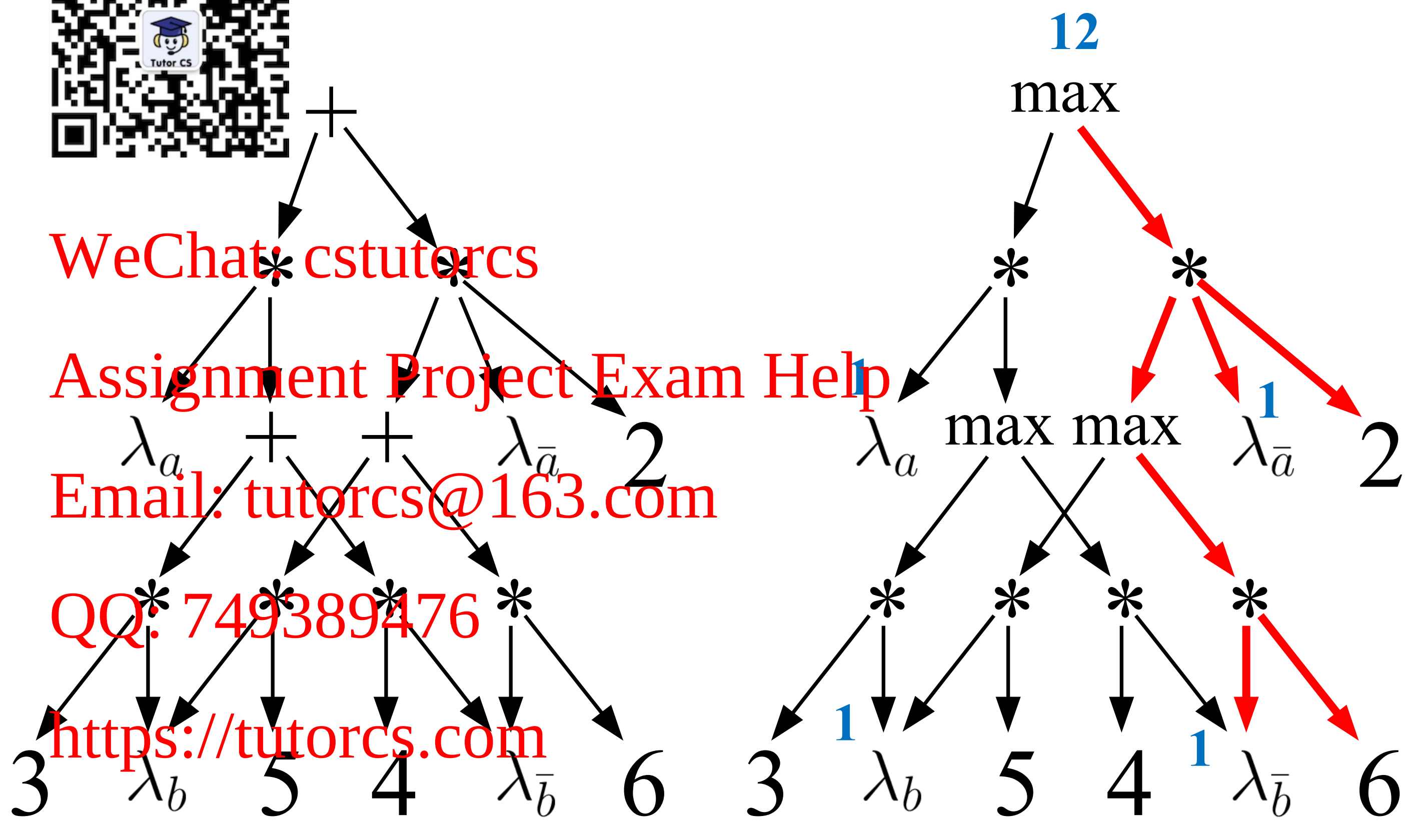
MPE Algorithm

Proposed in 2006, analysis in 2017

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12

term: $\bar{a}\,\bar{b}$

coefficient: 6*2 =12



finds largest coefficient of subcircuits compatible with evidence

Decomposable and Smooth ACs

Theorem 1. If an AC computes a factor f, and AC is decomposable and smooth, then the AC computes the marginals of f.

These circuits are known as **SPNs** (**Sum-Product Networks**, 2011)

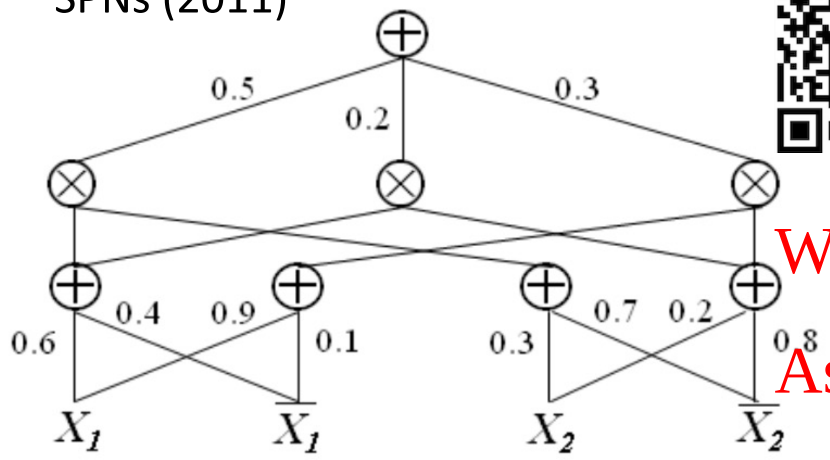
Deterministic, Decomposable and Smooth ACs

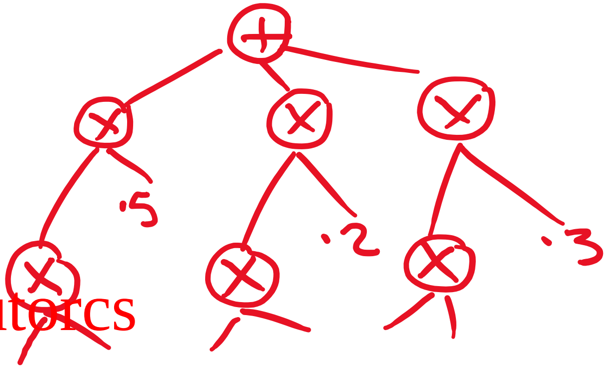
Theorem 3. If an AC is smooth, decomposable and deterministic, then to **compute MPE**, we can evaluate the maximizer circuit.

These are the original **ACs** (**Arithmetic Circuits**, 2003)

SPNs and Selective SPNs

SPNs (2011)





Arithmetic Circuit (AC)

Original (wrong) claim: SPNs also compute MPE in linear time

Later claim: SPNs with determinism compute MPE in linear time
These were called **selective SPNs**, which is another name for ACs that are decomposable, deterministic and smooth.

Why Determinism is Needed

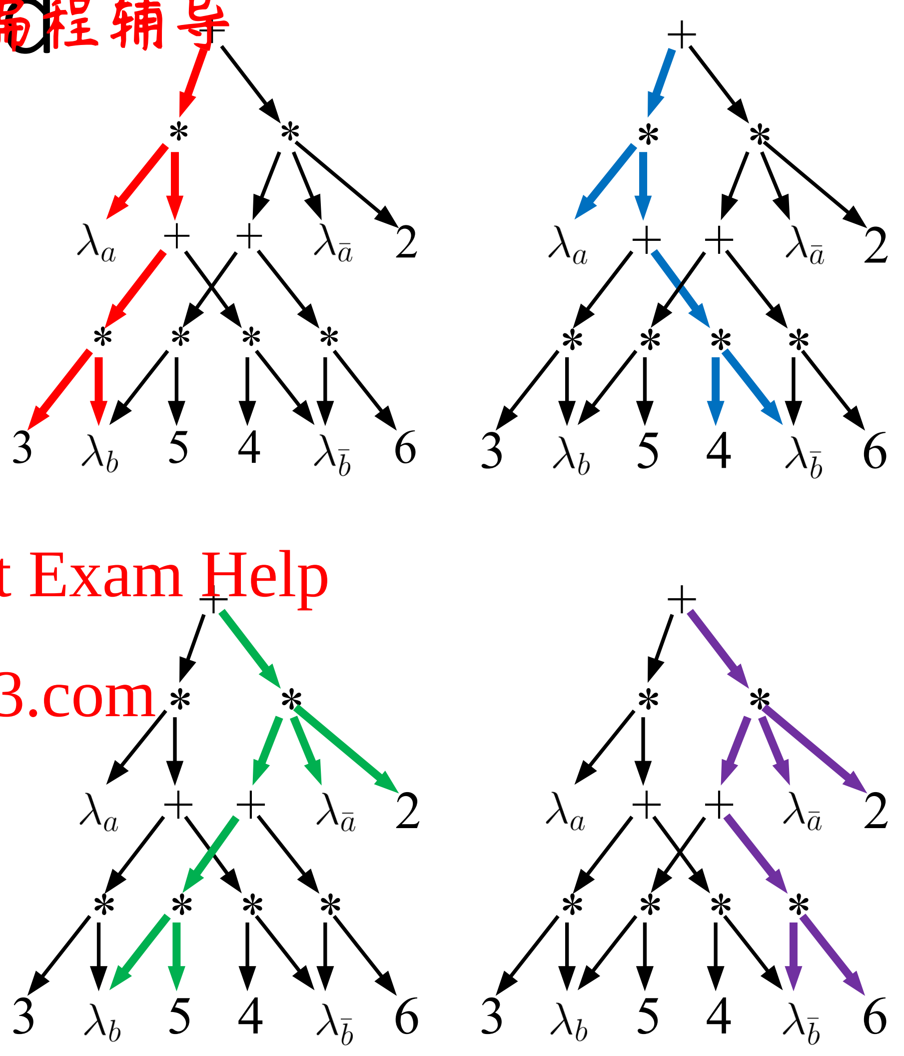
Decomposable & Smooth

subcircuit term is a complete variab[illegible]tion

Determinism →

1-to-1 mapping between subcircuits and complete variable instantiations

A	B	$f(A,B)$
a	b	3
a	$\bar{b}$	4
$\bar{a}$	b	10
$\bar{a}$	$\bar{b}$	12



excluding subcircuits with zero coefficients

Hardness of MPE on SPNs

Theorem 2. The MPE decision problem is NP-complete in decomposable and smooth ACs.

Original proof 2016, new proof 2017

Implications of Relaxing Determinism

- **Positive:** Separation
- **Negative:** Parametric Incompleteness

Separation

Theorem 4. There is a [illegible] of factors $f(X_1,\ldots, X_n)$ such that

(1) $f(X_1,\ldots, X_n)$ can be computed by a polysize decomposable and smooth AC

(2) any AC that computes $f(X_1,\ldots, X_n)$ and is deterministic, decomposable and smooth must have exponential size

Implications of Relaxing Determinism

- **Positive:** Separation
- **Negative:** Parametric Incompleteness

Complete Parameters

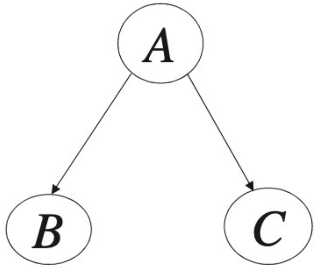
Definition 10 *A set of paramet*[illegible]*omplete for factor $f(\mathbf{X})$ iff for every instantiation* [illegible] *rameter $f(\mathbf{x})$ can be expressed as a product of parameters in Θ.*

$a_1b_1c_1$	.001
$a_1b_1c_2$	.009
$a_1b_2c_1$	.009
$a_1b_2c_2$	.081
$a_2b_1c_1$	.036
$a_2b_1c_2$	.144
$a_2b_2c_1$	.144
$a_2b_2c_2$	.576

$$\Theta = \{ .1,\ .2,\ .8,\ .9 \}$$

Complete Parameters

The parameters of a Bayesian network e for its factor



A	Θ_A
a_1	.1
a_2	.9

A	B	$\Theta_{B\|A}$
a_1	b_1	.1
a_1	b_2	.9
a_2	b_1	.2
a_2	b_2	.8

A	C	$\Theta_{C\|A}$
a_1	c_1	.1
a_1	c_2	.9
a_2	c_1	.2
a_2	c_2	.8

Network instantiation	
$a_1b_1c_1$	.001
$a_1b_1c_2$	.009
$a_1b_2c_1$	.009
$a_1b_2c_2$	.081
$a_2b_1c_1$	.036
$a_2b_1c_2$	.144
$a_2b_2c_1$	.144
$a_2b_2c_2$	.576

$$\Theta = \{\ .1,\ .2,\ .8,\ .9\ \}$$

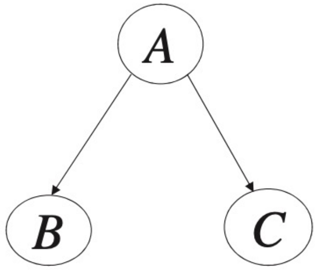
Notation

$$\mathcal{AC}(\mathbf{X}, \Theta)$$

Arithmetic Circuit (AC) over variables $\mathbf{X}$
and whose constants come from parameters Θ

Parametric Incompleteness

Theorem 5 (Completeness) *C[illegible] factor $f(\mathbf{X})$ and complete parameters Θ. There [illegible] an arithmetic circuit $\mathcal{AC}(\mathbf{X}, \Theta)$ that computes the factor marginals and is deterministic, decomposable and smooth.*



A	Θ_A
a_1	.1
a_2	.9

A	B	$\Theta_{B\|A}$
a_1	b_1	.1
a_1	b_2	.9
a_2	b_1	.2
a_2	b_2	.8

A	C	$\Theta_{C\|A}$
a_1	c_1	.1
a_1	c_2	.9
a_2	c_1	.2
a_2	c_2	.8

$$\Theta = \{ .1,\ .2,\ .8,\ .9 \}$$

Theorem 5 does not hold if we drop determinism

Enforcing decomposability and smoothness is not sufficient: **You need to search for AC parameters**

Parametric Incompleteness

Boolean Factor

A	B	C	
T	T	T	0
T	T	F	1
T	F	T	1
T	F	F	1
F	T	T	0
F	T	F	1
F	F	T	0
F	F	F	0

An AC that computes marginals of a Boolean factor is an AC that can count models

$$\Theta = \{0, 1\}$$

Theorem 6 (Parametric Incompleteness) *Let $f(\mathbf{X})$ be a Boolean factor and $\Theta = \{0, 1\}$ (Θ is complete for f). A circuit $\mathcal{AC}(\mathbf{X}, \Theta)$ cannot compute $f(\mathbf{X})$ if it is decomposable, smooth and free of dead nodes, but not deterministic.*

ACs, SPNs, and Probabilistic Circuits

- Loose usage:
  Computation graphs, with indicator variab[illegible], adders and multipliers
- Computer theory: ACs (not t[illegible]s)
- Property-driven (more precise usage):
  - **AC (2003)**: decomposable, deterministic and smooth
  - **SPN (2011)**: decomposable and smooth
  - **Selective SPN**: decomposable, deterministic and smooth (= AC)
- Probabilistic Circuits (umbrella term):
  Restrictive as factors do not need to represent probabilities (they are general mappings from instantiations to positive numbers)
- There are other circuit types:
  - **PSDD (2014)**: structured decomposability, partitioned determinism
  - ….
- Guide: Focus on properties
  All implications of what we studied/know on the Boolean side

Big Ideas & Milestones

A Differential Approach to Inference in [illegible]etworks
Adnan Darwiche, JACM 2003

Introduced ACs (decomposable, smooth, deterministic)

On the Robustness of Most Probable Explanations
Hei Chan, Adnan Darwiche, UAI 2006

Introduced MPE algorithm for ACs (and subcircuits)

Learning Arithmetic Circuits
Daniel Lowd, Pedro Domingos, UAI 2008

Introduced learning ACs from data

Sum-Product Networks
Hoifung Poon, Pedro Domingos, UAI 2011

Introduced handcrafting ACs
Dropped determinism

On Relaxing Determinism in Arithmetic Circuits
Arthur Choi and Adnan Darwiche, ICML 2017

Provided a comprehensive theory that explained roles of decomposability, smoothness and determinism (this lecture)

Tractable Boolean and Arithmetic Circuits

Adnan Darwiche, Computer Sc[illegible]rtment, UCLA

Tractable Boolean and arithmetic circuits have been studied extensively in AI for over two decades now. These circuits were initially proposed as "compiled objects," meant to facilitate logical and probabilistic reasoning, as they permit various types of inference to be performed in linear-time and a feed-forward fashion like neural networks. In more recent years, the role of tractable circuits has significantly expanded as they became a computational and semantical backbone for some approaches that aim to integrate knowledge, reasoning and learning. In this chapter, we review the foundations of tractable circuits and some associated milestones, while focusing on their core properties and techniques that make them particularly useful for the broad aims of neuro-symbolic AI.

Book on Neuro-Symbolic AI (2022): being integrated into ML pipelines