Exploiting Functional Dependence in Bayesian Network Inference

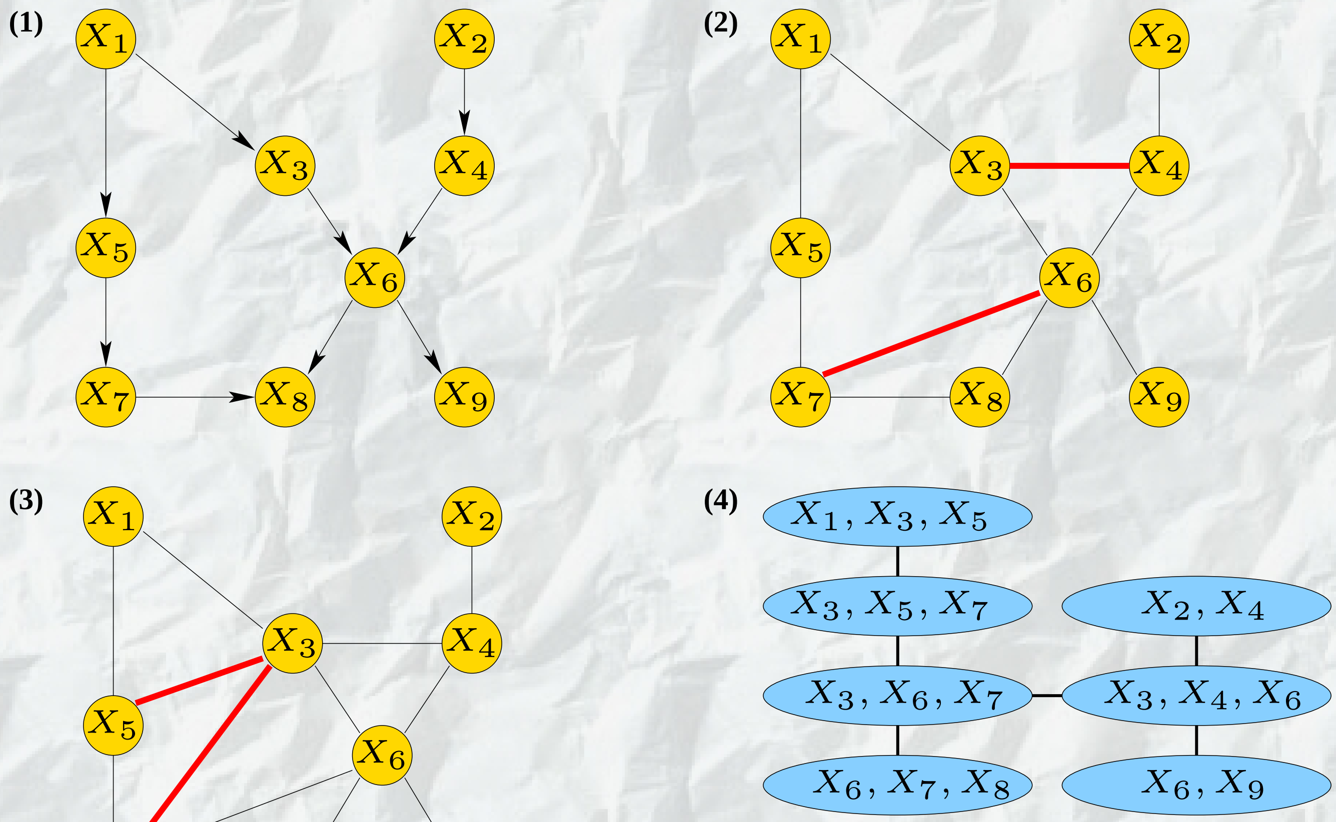
Jiří Vomlel

Laboratory for Intelligent Systems
University of Economics, Prague

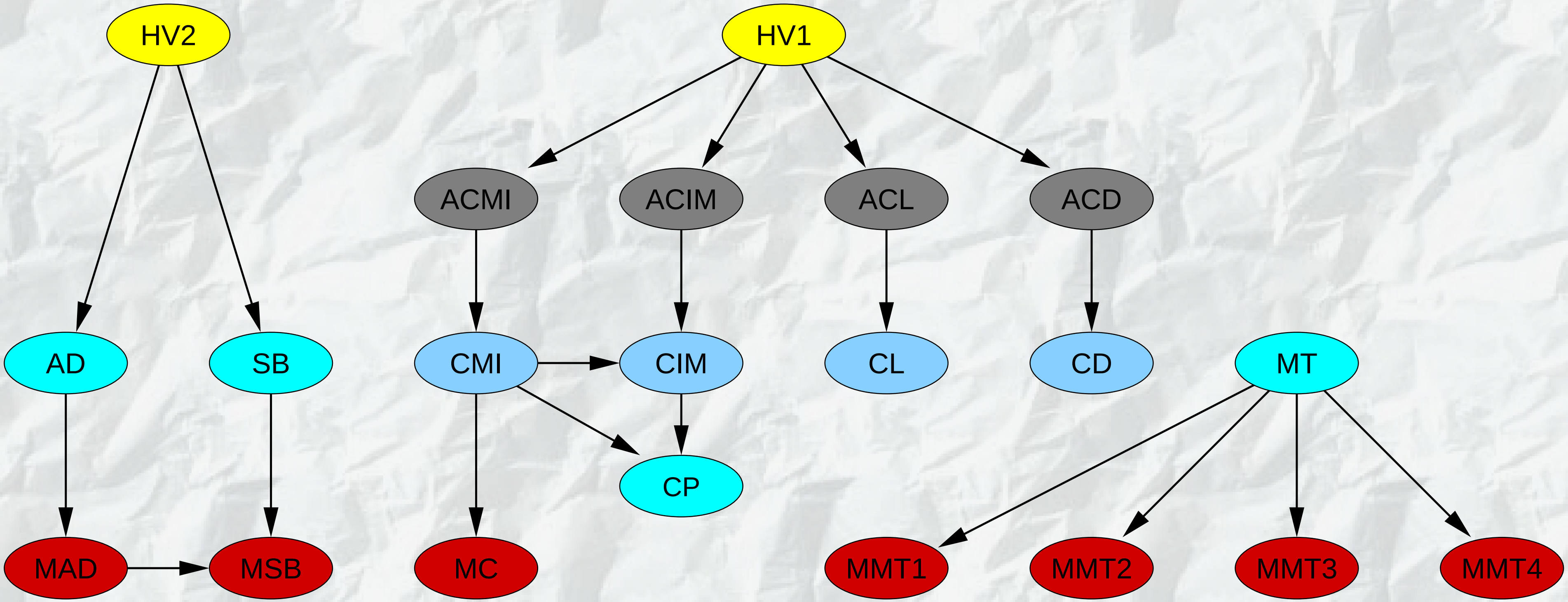
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic

http://www.utia.cas.cz/vomlel/

Bayesian networks: junction tree propagation



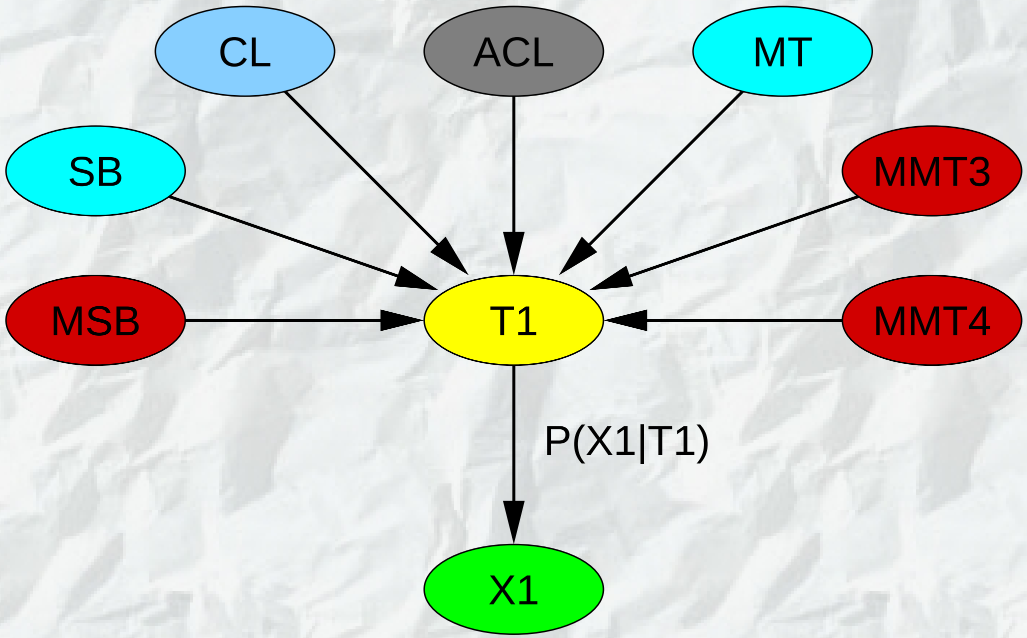
Original model of a student



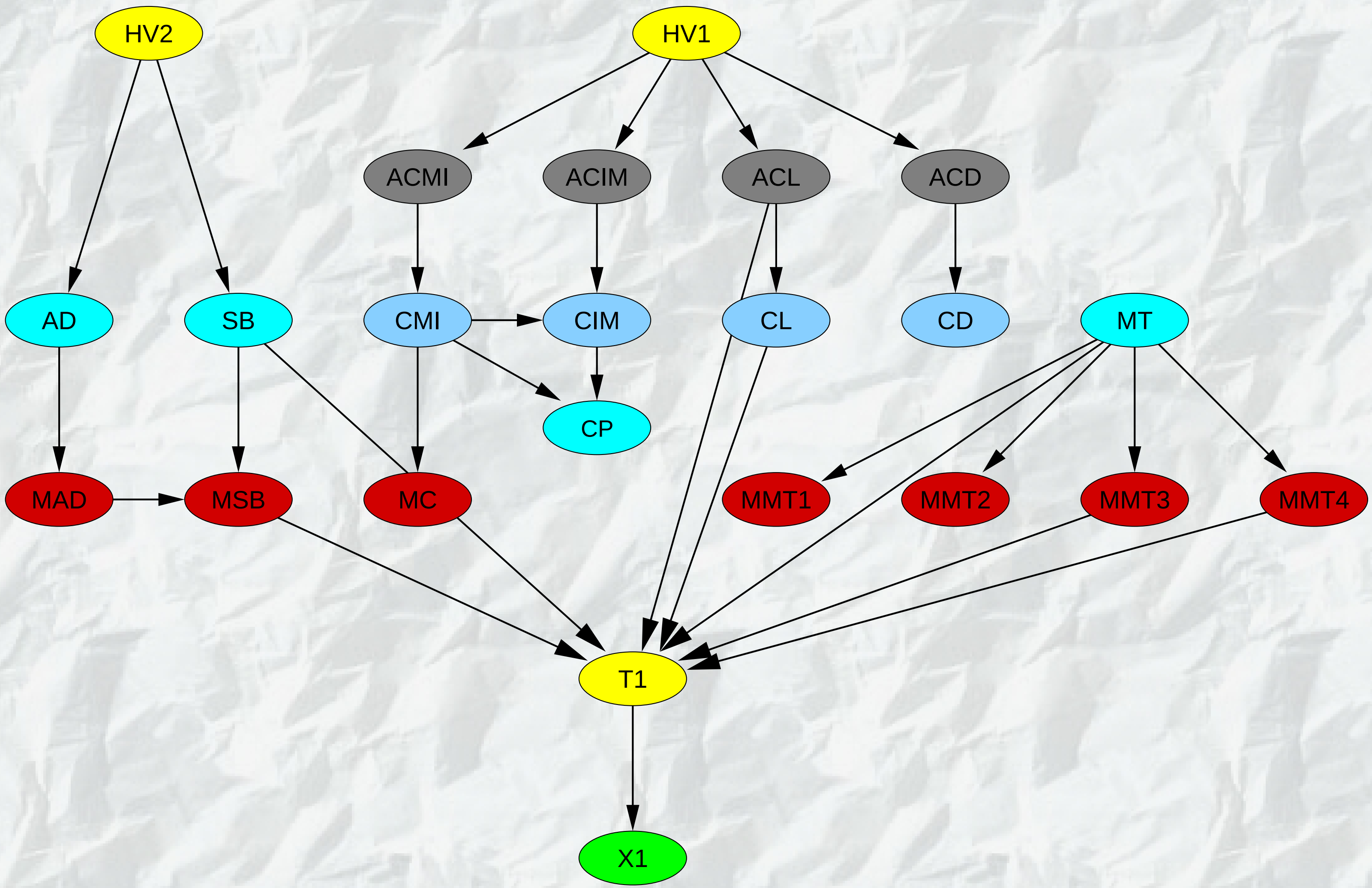
Evidence model of a task

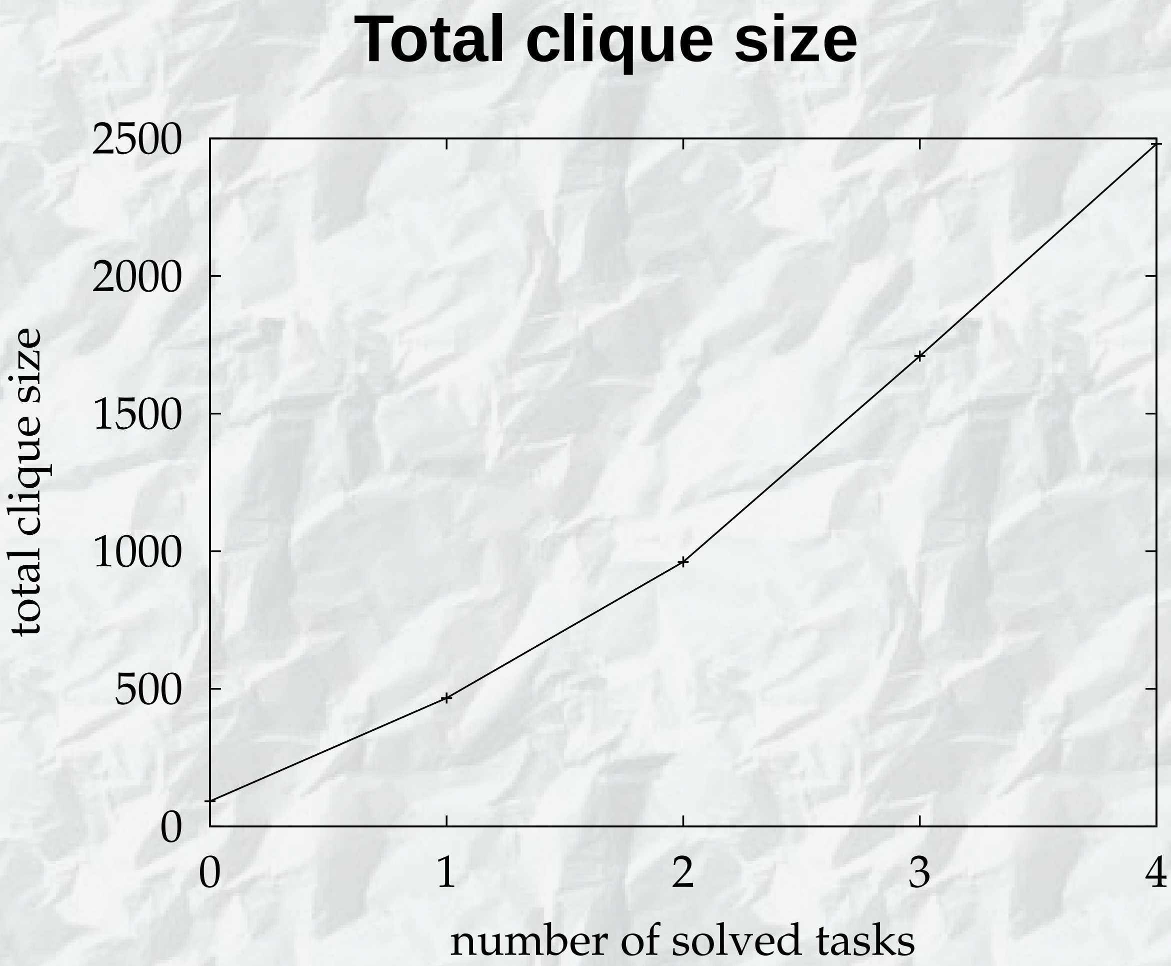
$$\left(\frac{3}{4}\cdot\frac{5}{6}\right)-\frac{1}{8}=\frac{15}{24}-\frac{1}{8}=\frac{5}{8}-\frac{1}{8}=\frac{4}{8}=\frac{1}{2}$$

$$T_1 \quad \Leftrightarrow \quad MT \,\&\, CL \,\&\, ACL \,\&\, SB \,\&\, \neg MMT3 \,\&\, \neg MMT4 \,\&\, \neg MSB$$



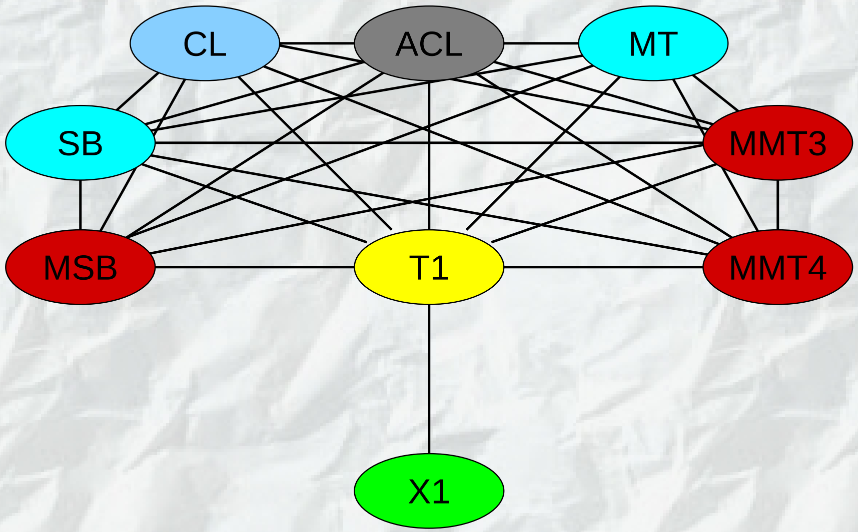
Original model + one evidence model



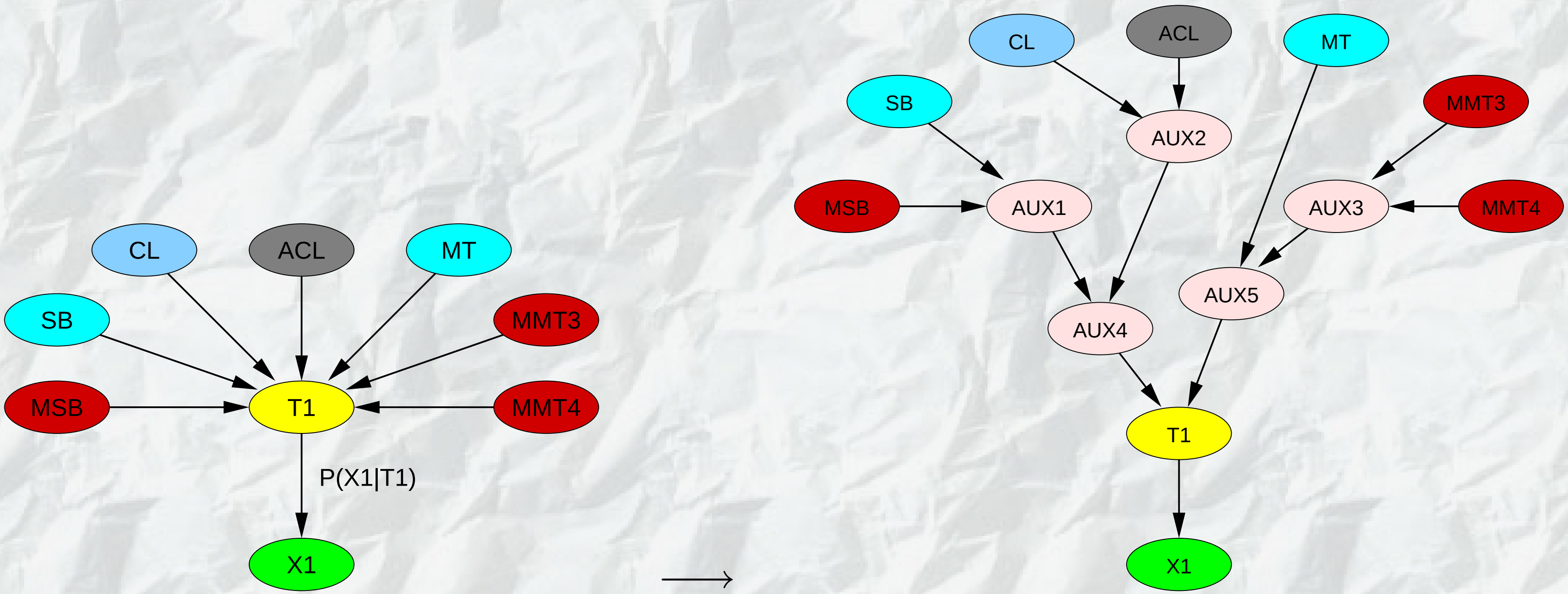

Total clique size
total clique size
number of solved tasks
0
500
1000
1500
2000
2500
0
1
2
3
4

This is the problem ...

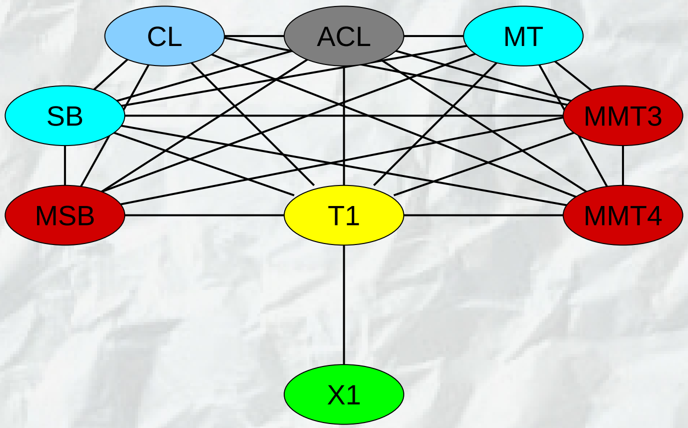
... all parents get inter-connected.



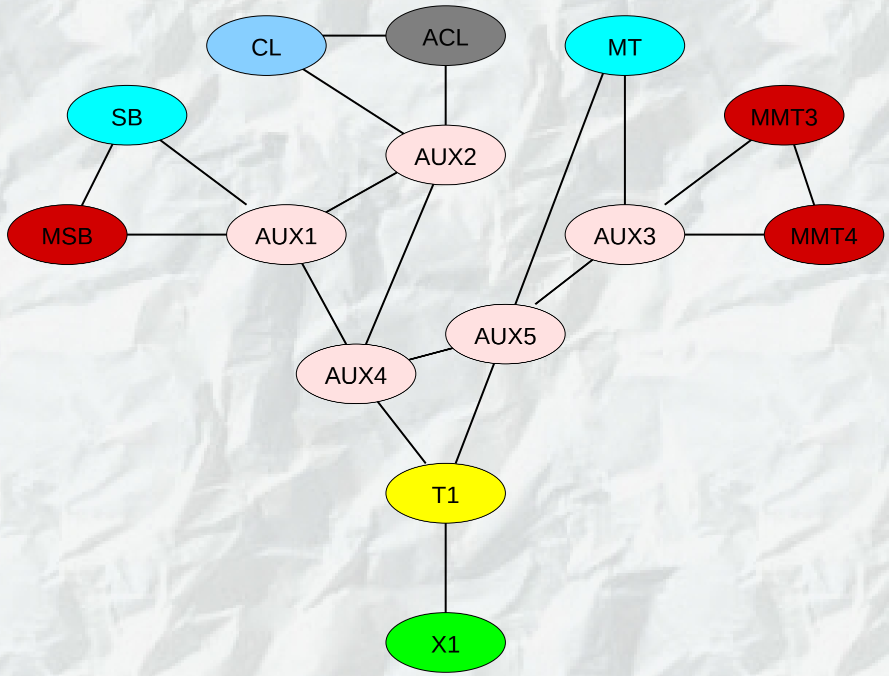
Parent divorcing (Olesen et al., 1989)

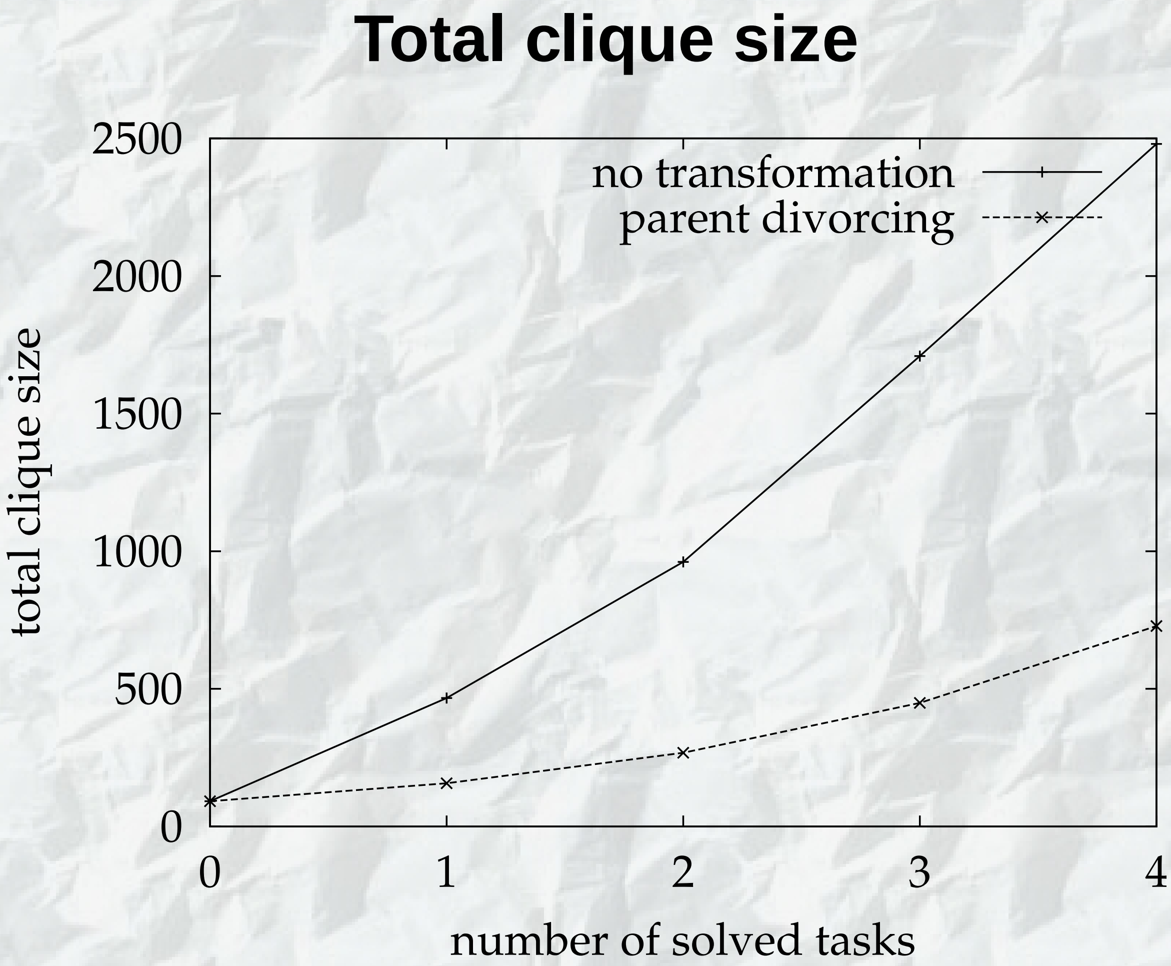


Parent divorcing



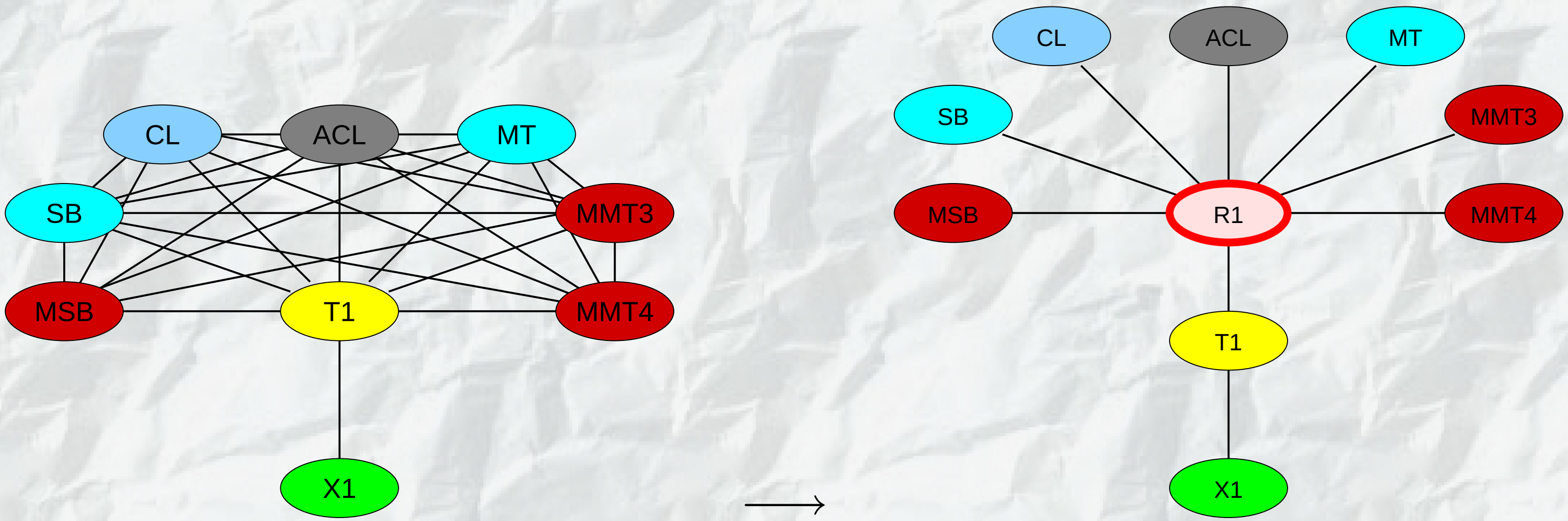
⟶



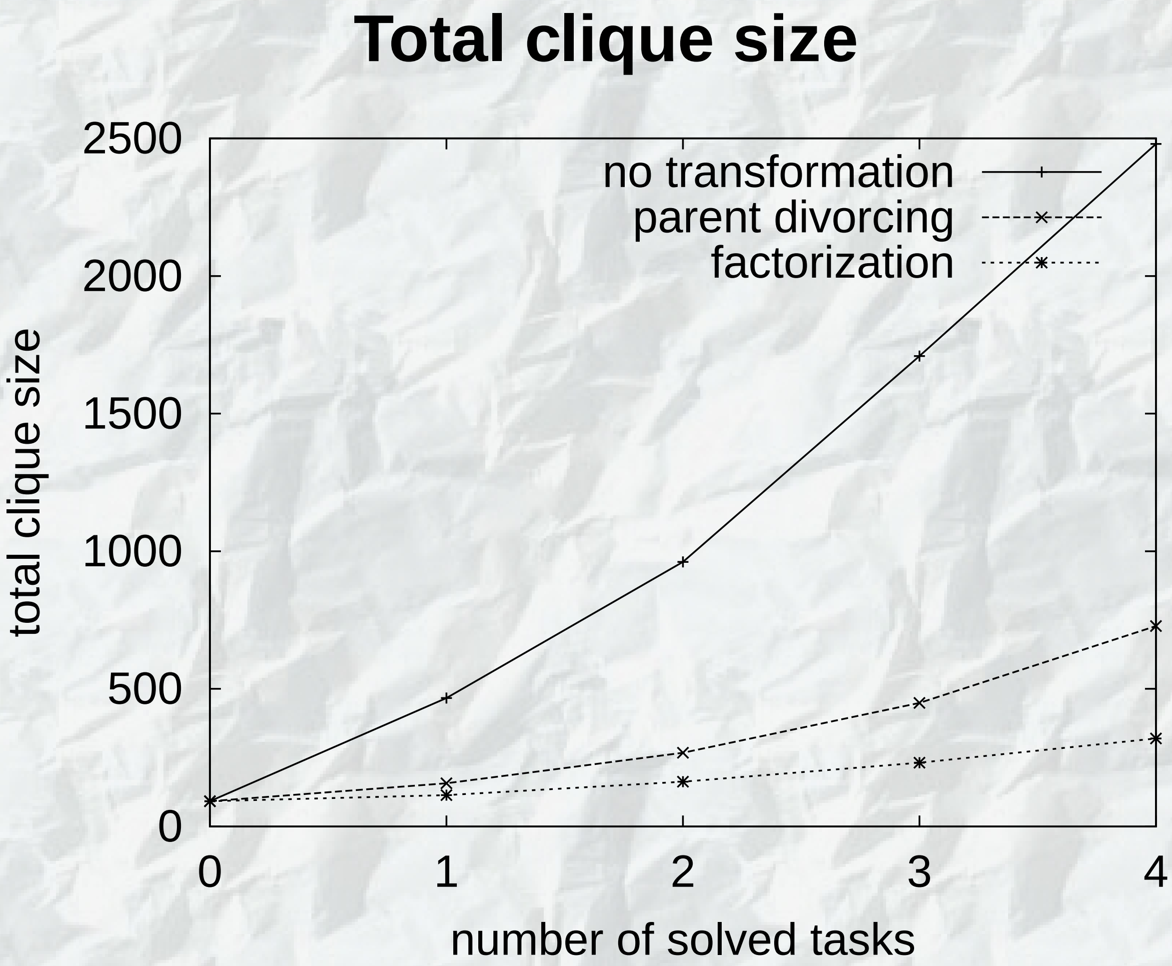


Total clique size
total clique size
2500
2000
1500
1000
500
0
0
1
2
3
4
number of solved tasks
no transformation
parent divorcing

Factorized evidence model



$$\psi(T1, MSB, SB, CL, ACL, MT, MMT3, MMT4) = \sum_{R1} \left(\begin{array}{l} \varphi(T1, R1) \cdot \varphi(R1, MSB) \cdot \varphi(R1, SB) \cdot \varphi(R1, CL) \cdot \\ \varphi(R1, ACL) \cdot \varphi(R1, MT) \cdot \varphi(R1, MMT3) \cdot \varphi(R1, MMT4) \end{array} \right)$$

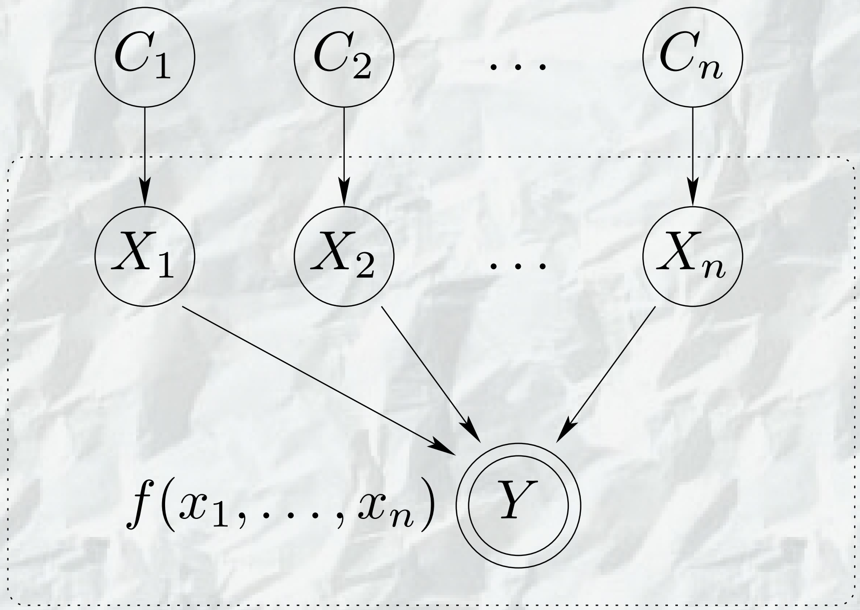

Total clique size
no transformation
parent divorcing
factorization
total clique size
0
500
1000
1500
2000
2500
0
1
2
3
4
number of solved tasks

Functional dependence

Y is functionally dependent on $X_1, \ldots, X_n$ if

$$P(Y = y, X_1 = x_1, \ldots, X_n = x_n) = \begin{cases} 1 & \text{if } y = f(x_1, \ldots, x_n) \\ 0 & \text{otherwise.} \end{cases}$$

Example - a model with independence of causal influence



Factorization of MAX

$$y = f(x_1, x_2) \; = \; \max\,\{x_1, x_2\}$$

$$P(Y|X_1, X_2) = \sum_R h(Y, R) \cdot g_1(X_1, R) \cdot g_2(X_2, R)$$

	0			+1			+2		
	0	+1	+2	0	+1	+2	0	+1	+2
0	1								
+1		1		1	1				
+2			1			1	1	1	1

$=$

$\sum_{r_1, r_2, r_3}($

	r_1	r_2	r_3
0	1		
+1	-1	1	
+2		-1	1

$\times$

	r_1	r_2	r_3
0	1	1	1
+1		1	1
+2			1

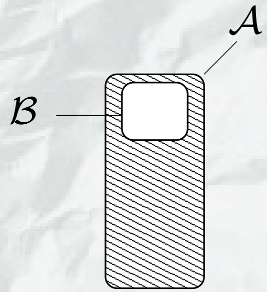
$\times$

	r_1	r_2	r_3
0	1	1	1
+1		1	1
+2			1

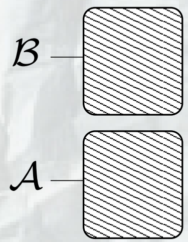
$)$

Proper difference and disjunctive union

- If $\mathcal{A} \supseteq \mathcal{B}$ then proper difference of $\mathcal{A}$ and $\mathcal{B}$ is defined as $\mathcal{A} \ominus \mathcal{B} = \{\mathbf{x} \in \mathcal{A} \wedge \mathbf{x} \notin \mathcal{B}\}$.



- If $\mathcal{A} \cap \mathcal{B} = \emptyset$ then disjunctive union of $\mathcal{A}$ and $\mathcal{B}$ is defined as $\mathcal{A} \oplus \mathcal{B} = \{\mathbf{x} \in \mathcal{A} \vee \mathbf{x} \in \mathcal{B}\}$.

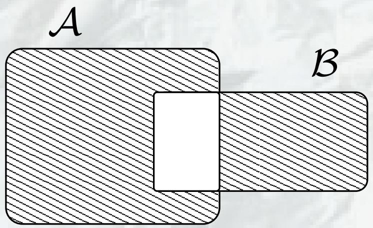


Special case of symmetric difference

If $\mathcal{A} \supseteq \mathcal{B}$ or $\mathcal{A} \cap \mathcal{B} = \emptyset$ then previous two operations can be replaced by a single operation of *symetric difference* of $\mathcal{A}$ and $\mathcal{B}$ defined as

$$\{\mathbf{x} \in \mathcal{A} \wedge \mathbf{x} \notin \mathcal{B}\} \ \cup \ \{\mathbf{x} \in \mathcal{B} \wedge \mathbf{x} \notin \mathcal{A}\} \ .$$

Example when symmetric difference can not be applied:

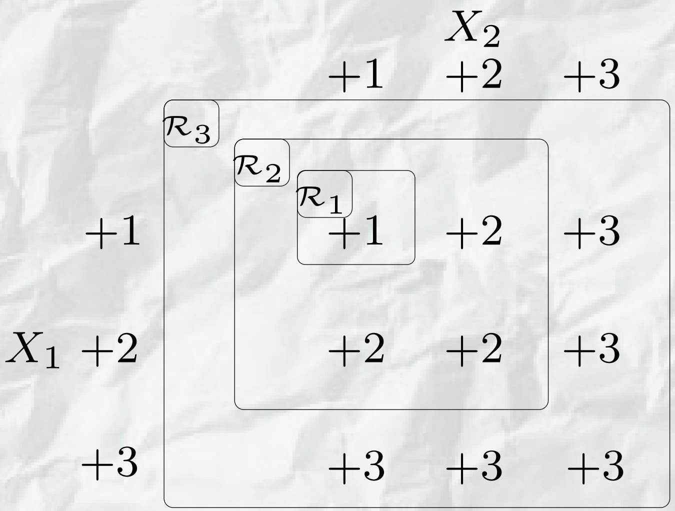


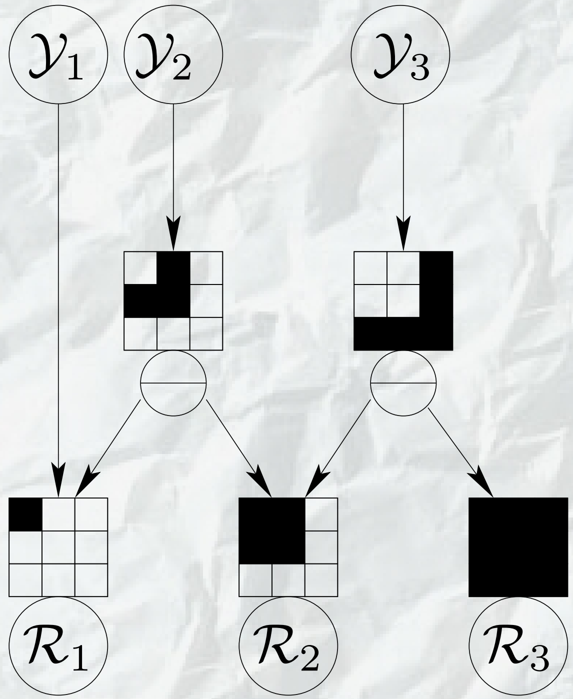
Minimal generator of rectangles (MGR)

For a given partition $\mathcal{Y}_1, \dots, \mathcal{Y}_q$ of $\mathcal{X} = \times_{i=1}^{n} \mathcal{X}_i$
find a set $\{\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_k\}$ of minimal cardinality such that:

- for $j = 1, \dots, k$ set $\mathcal{R}_j$ is a rectangle,
  i.e. $\mathcal{R}_j = \times_{i=1}^{n} \mathcal{D}_i$, $\emptyset \neq \mathcal{D}_i \subseteq \mathcal{X}_i$,
- each element $\mathcal{Y}_\ell, \ell = 1, 2, \dots q$ of the partition can be generated from the generator $\{\mathcal{R}_1, \dots, \mathcal{R}_k\}$ using operations $\ominus$ and $\oplus$.

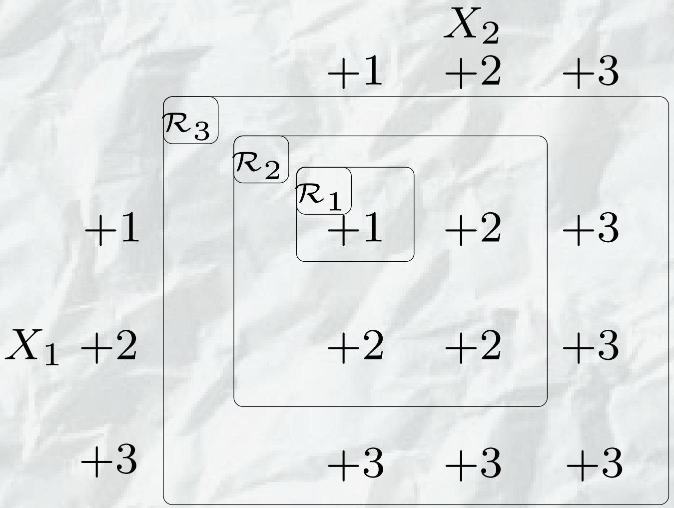
Factorization of MAX: problem reformulation





$$
\begin{aligned}
\mathcal{Y}_1 &= \{\ (+1,+1)\ \} \quad = \mathcal{R}_1 \\
\mathcal{Y}_2 &= \{\ (+1,+2),(+2,+1),(+2,+2)\ \} \quad = \mathcal{R}_2 \ominus \mathcal{R}_1 \\
\mathcal{Y}_3 &= \{\ (+1,+3),(+2,+3),(+3,+1),(+3,+2)(+3,+3\}) \quad = \ \mathcal{R}_3 \ominus \mathcal{R}_2
\end{aligned}
$$

Correspondence of MGR and factorization for MAX



$$\mathcal{Y}_1 = \mathcal{R}_1$$
$$\mathcal{Y}_2 = \mathcal{R}_2 \ominus \mathcal{R}_1$$
$$\mathcal{Y}_3 = \mathcal{R}_3 \ominus \mathcal{R}_2$$

Hidden variable R has one state for each rectangle

$h(Y, R)$

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$
$+1$	$+1$	0	0
$+2$	-1	$+1$	0
$+3$	0	-1	$+1$

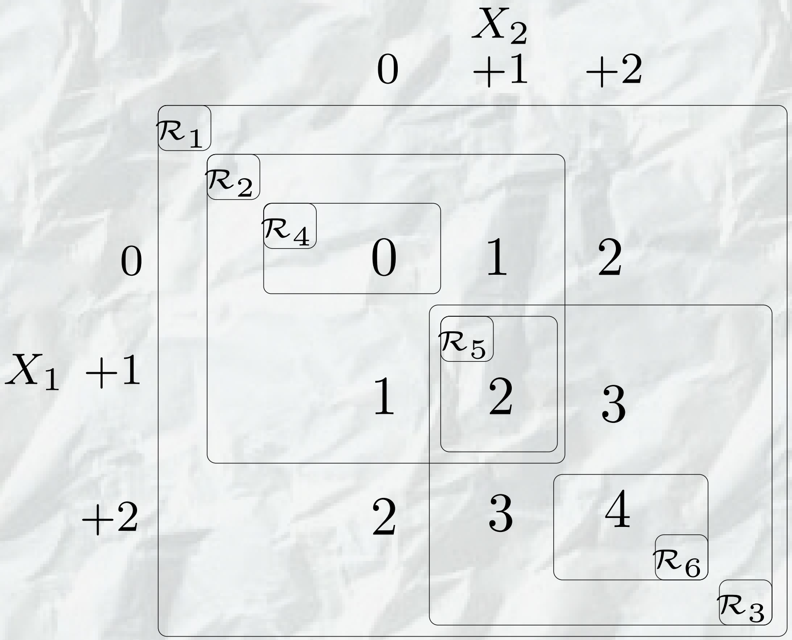
$g_1(X_1, R)$

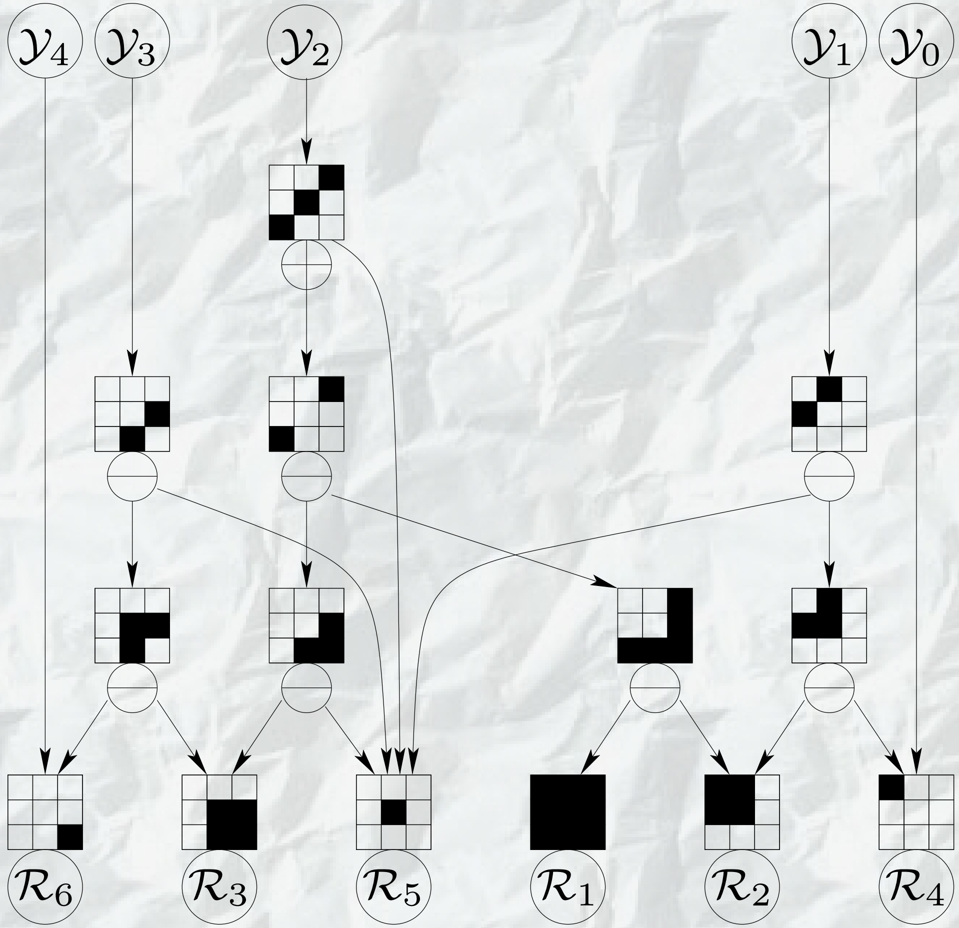
	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$
$+1$	1	1	1
$+2$	0	1	1
$+3$	0	0	1

$g_2(X_2, R)$

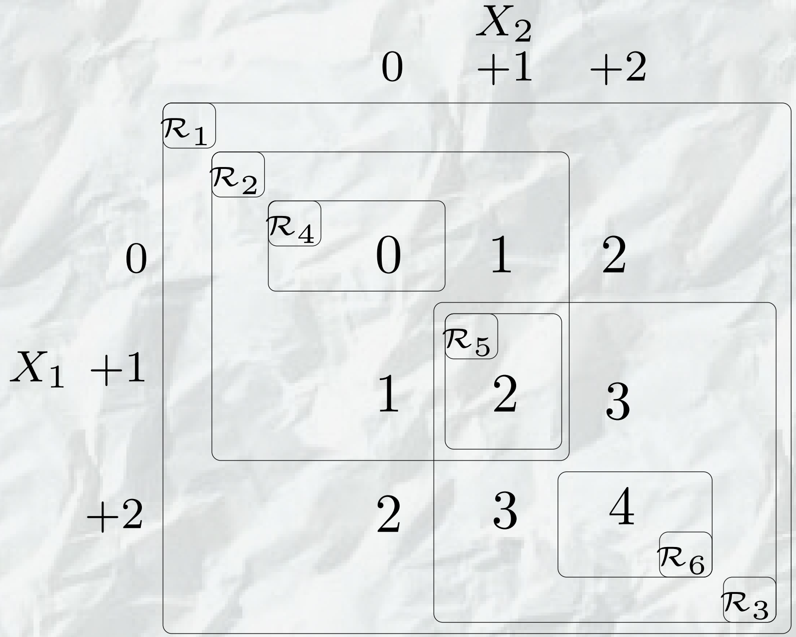
	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$
$+1$	1	1	1
$+2$	0	1	1
$+3$	0	0	1

Minimal generator of rectangles for ADD





Correspondence of MGR and factorization for ADD



$$
\begin{aligned}
\mathcal{Y}_1 &= \mathcal{R}_4 \\
\mathcal{Y}_2 &= \mathcal{R}_2 \ominus \mathcal{R}_4 \ominus \mathcal{R}_5 \\
\mathcal{Y}_3 &= (\mathcal{R}_1 \ominus (\mathcal{R}_2 \ominus \mathcal{R}_5)) \ominus (\mathcal{R}_3 \ominus \mathcal{R}_5) \\
\mathcal{Y}_4 &= \mathcal{R}_3 \ominus \mathcal{R}_5 \ominus \mathcal{R}_6 \\
\mathcal{Y}_5 &= \mathcal{R}_6 \ ,
\end{aligned}
$$

Hidden variable B has one state for each rectangle

$h(Y, R)$

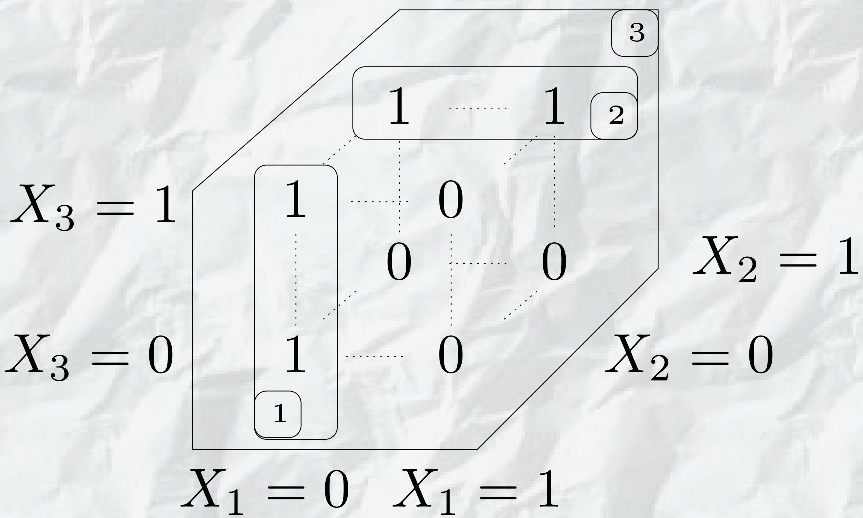
	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_4$	$\mathcal{R}_5$	$\mathcal{R}_6$
0	0	0	0	+1	0	0
+1	0	+1	0	−1	−1	0
+2	+1	−1	−1	0	+2	0
+3	0	0	+1	0	−1	−1
+4	0	0	0	0	0	+1

$g_1(X_1, R), \quad g_2(X_2, R)$

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_4$	$\mathcal{R}_5$	$\mathcal{R}_6$
0	1	1	0	1	0	0
+1	1	1	1	0	1	0
+2	1	0	1	0	0	1

A Boolean function

$$
\begin{aligned}
Y &= (X_1 \vee X_2) \Rightarrow (X_2 \wedge X_3) \\
&= (\neg X_1 \wedge \neg X_2) \vee (X_2 \wedge X_3)
\end{aligned}
$$



$$
\begin{aligned}
\mathcal{Y}_0 &= \mathcal{R}_3 \ominus (\mathcal{R}_2 \oplus \mathcal{R}_1) \\
\mathcal{Y}_1 &= \mathcal{R}_2 \oplus \mathcal{R}_1
\end{aligned}
$$

Open questions and problems

- A general algorithm that for a given function finds its minimal generator.
- Generalization of the method so that tables $g_1(X_1, R), \ldots, g_n(X_n, R)$ need not contain zeroes and ones only
- A method for approximation of any probability table by a product of two-dimensional tables.
- Links to communication complexity.
- ...

...and it is time to end.