COMPOSITIONAL MODELS

COMPOSITIONAL MODELS

Oligodimensional distributions

$$P_1((x_i)_{i \in K_1}), P_2((x_i)_{i \in K_2}), \ldots, P_n((x_i)_{i \in K_n})$$

COMPOSITIONAL MODELS

Oligodimensional distributions

$$P_1((x_i)_{i\in K_1}), P_2((x_i)_{i\in K_2}), \ldots, P_n((x_i)_{i\in K_n})$$

$$P_1(x_{K_1}), P_2(x_{K_2}), \ldots, P_n(x_{K_n})$$

Definition of the operator of composition

For 2 distributions:

$P(x_K), Q(x_L)$

their *composition* is defined by the expression

$$P \triangleright Q = \frac{PQ}{Q(x_{K \cap L})}$$

Definition of the operator of composition

For 2 distributions:

$P(x_K), Q(x_L)$

their *composition* is defined by the expression

$$P \triangleright Q = \frac{PQ}{Q(x_{K \cap L})} = P(x_K)Q(x_{L \setminus K}|x_{K \cap L})$$

Definition of the operator of composition

For 2 distributions:
$P(x_K), Q(x_L)$
their *composition* is defined by the expression

$$P \triangleright Q = \frac{PQ}{Q(x_{K \cap L})} = P(x_K)Q(x_{L \setminus K}|x_{K \cap L})$$

More precisely:

$$P \triangleright Q = \begin{cases} \frac{P(x_K)Q(x_L)}{Q(x_{K\cap L})} & \text{if } P(x_{K \cap L}) \ll Q(x_{K \cap L}), \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Definition of the "left composition"

For 2 distributions:
$P(x_K), Q(x_L)$

$$P \triangleleft Q = \begin{cases} \dfrac{P(x_K)Q(x_L)}{P(x_{K\cap L})} & \text{if } Q(x_{K\cap L}) \ll P(x_{K\cap L}), \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Basic properties of the operators of composition

If $P \rhd Q$ and $P \lhd Q$ are defined then

Basic properties of the operators of composition

If $P \triangleright Q$ and $P \triangleleft Q$ are defined then

- $P \triangleright Q$ is a distribution of variables $X_{K \cup L}$

Basic properties of the operators of composition

If $P \triangleright Q$ and $P \triangleleft Q$ are defined then

- $P \triangleright Q$ is a distribution of variables $X_{K \cup L}$
- P is a marginal distribution of $P \triangleright Q$
  Q is a marginal distribution of $P \triangleleft Q$

Basic properties of the operators of composition

If $P \triangleright Q$ and $P \triangleleft Q$ are defined then

- $P \triangleright Q$ is a distribution of variables $X_{K \cup L}$
- P is a marginal distribution of $P \triangleright Q$

  Q is a marginal distribution of $P \triangleleft Q$
- $P \triangleright Q = P \triangleleft Q$ *iff* P and Q are consistent:

  tj. $P(x_{K \cap L}) = Q(x_{K \cap L})$

Basic properties of the operators of composition

If $P \triangleright Q$ and $P \triangleleft Q$ are defined then

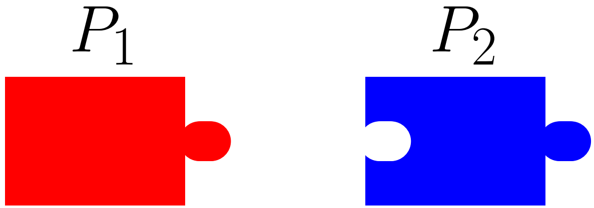
- $P \triangleright Q$ is a distribution of variables $X_{K \cup L}$
- P is a marginal distribution of $P \triangleright Q$
  Q is a marginal distribution of $P \triangleleft Q$
- $P \triangleright Q = P \triangleleft Q$ *iff* P and Q are consistent:
  tj. $P(x_{K \cap L}) = Q(x_{K \cap L})$
- $X_{K \setminus L} \perp\!\!\!\perp X_{L \setminus K} | X_{K \cap L} \; [P \triangleright Q]$

Basic properties of the operators of composition

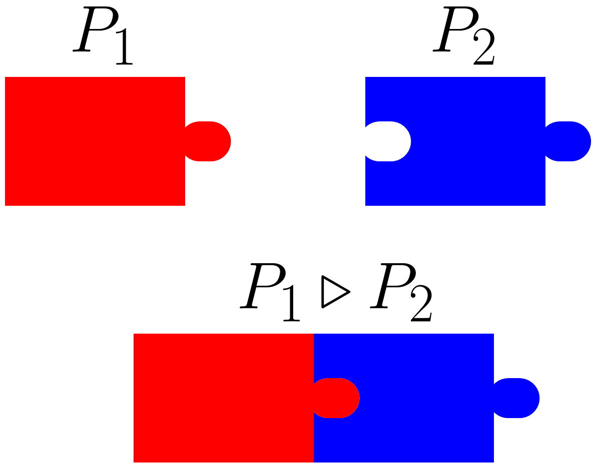
If $P \triangleright Q$ and $P \triangleleft Q$ are defined then

- $P \triangleright Q$ is a distribution of variables $X_{K \cup L}$
- P is a marginal distribution of $P \triangleright Q$
  Q is a marginal distribution of $P \triangleleft Q$
- $P \triangleright Q = P \triangleleft Q$ *iff* P and Q are consistent:
  tj. $P(x_{K \cap L}) = Q(x_{K \cap L})$
- $X_{K \setminus L} \perp\!\!\!\perp X_{L \setminus K} | X_{K \cap L} \; [P \triangleright Q]$
- if P and Q are consistent then $P \triangleright Q$ is a maximum entropy extension of both P and Q.

Composition of distributions



Composition of distributions



Generating sequence

System $\{P_j(x_{K_j}) \ : j = 1, 2, \ldots, m\}$.

Compositional model is defined by a *generating sequence* $(m = 2)$:

P_1, P_2 define distribution:

$P_1(x_{K_1}) \triangleright P_2(x_{K_2})$

which is defined for variables $X_{(K_1 \cup K_2)}$

Generating sequence

System $\{P_j(x_{K_j}) \; : j = 1, 2, \ldots, m\}$.

Compositional model is defined by a *generating sequence* $(m = 3)$:

P_1, P_2, P_3 define distribution:

$$P_1(x_{K_1}) \triangleright P_2(x_{K_2}) \triangleright P_3(x_{K_3})$$
$$= (P_1 \triangleright P_2) \triangleright P_3$$

which is defined for variables $X_{(K_1 \cup K_2 \cup K_3)}$

Generating sequence

System $\{P_j(x_{K_j}) : j = 1, 2, \ldots, m\}$.

Compositional model is defined by a *generating sequence* $(m = 4)$:

P_1, P_2, P_3, P_4 define distribution:

$$P_1(x_{K_1}) \triangleright P_2(x_{K_2}) \triangleright P_3(x_{K_3}) \triangleright P_4(x_{K_4})$$
$$= ((P_1 \triangleright P_2) \triangleright P_3) \triangleright P_4$$

which is defined for variables $X_{(K_1 \cup K_2 \cup K_3 \cup K_4)}$

Generating sequence

System $\{P_j(x_{K_j}) : j = 1, 2, \ldots, m\}$.

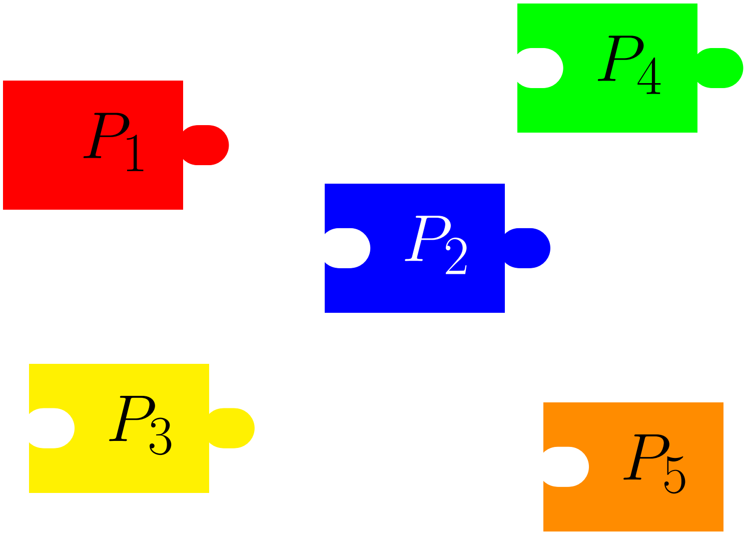
Compositional model is defined by a *generating sequence*:

$P_1, \ldots, P_m$ define distribution:

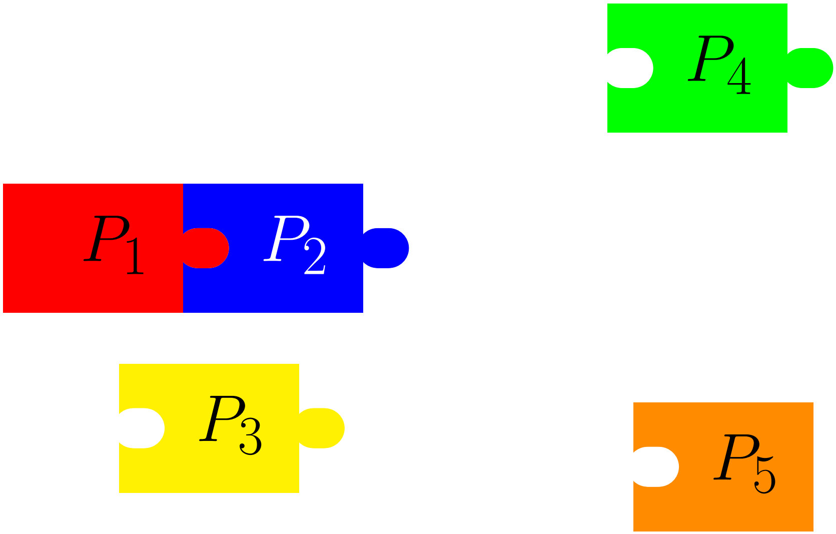
$$P_1(x_{K_1}) \triangleright P_2(x_{K_2}) \triangleright \ldots \triangleright P_m(x_{K_m})$$
$$= (\ldots(P_1 \triangleright P_2) \triangleright \ldots) \triangleright P_m)$$

which is defined for variables $X_{(K_1 \cup \ldots \cup K_m)}$

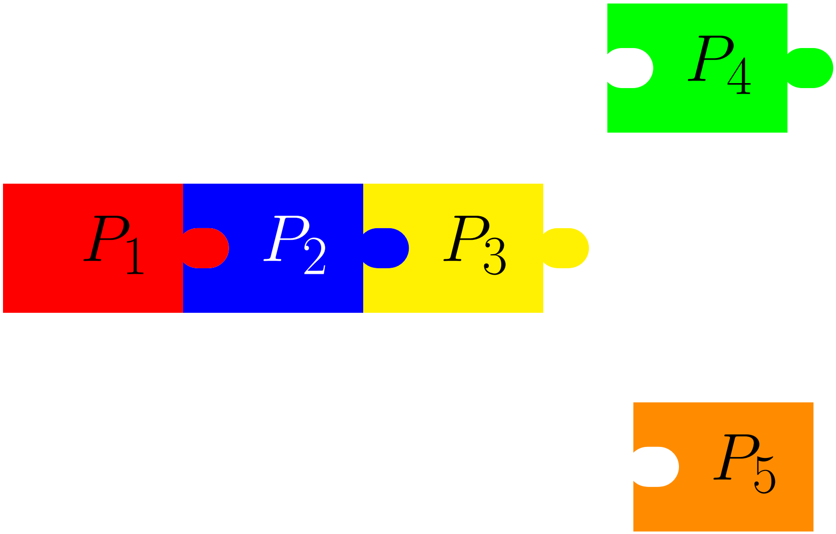
Composition of oligodimensional distributions



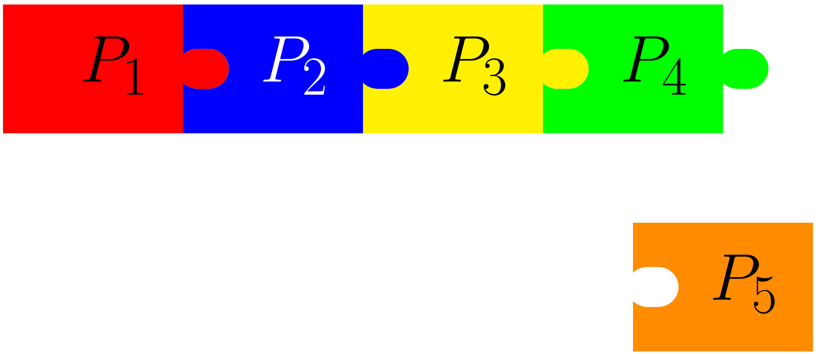
Composition of oligodimensional distributions



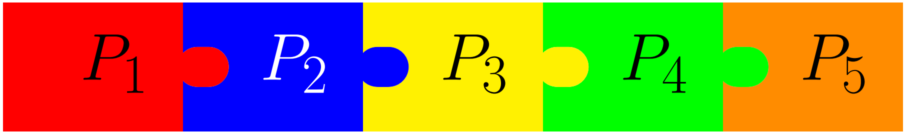
Composition of oligodimensional distributions



Composition of oligodimensional distributions



Composition of oligodimensional distributions



Why we introduced two operators of composition?

Why we introduced two operators of composition?

Operators are neither commutative nor associative - saving brackets; operators are applied from left to right.

Why we introduced two operators of composition?

Operators are neither commutative nor associative - saving brackets; operators are applied from left to right.

Main difference between them is, however,

Why we introduced two operators of composition?

Operators are neither commutative nor associative - saving brackets; operators are applied from left to right.

Main difference between them is, however, – their computational complexity.

Why we introduced two operators of composition?

Operators are neither commutative nor associative - saving brackets; operators are applied from left to right.

Main difference between them is, however, – their computational complexity.

Denote $L = K_j \cap (K_1 \cup \ldots \cup Kj - 1)$,

$$(P_1 \triangleright \ldots \triangleright P_{j-1}) \triangleright P_j = \frac{(P_1 \triangleright \ldots \triangleright P_{j-1}) \cdot P_j}{(P_j)^{\downarrow L}}$$

Why we introduced two operators of composition?

Operators are neither commutative nor associative - saving brackets; operators are applied from left to right.

Main difference between them is, however, – their computational complexity.

Denote $L = K_j \cap (K_1 \cup \ldots \cup Kj - 1)$,

$$(P_1 \triangleleft \ldots \triangleleft P_{j-1}) \triangleleft P_j = \frac{(P_1 \triangleleft \ldots \triangleleft P_{j-1}) \cdot P_j}{(P_1 \triangleleft \ldots \triangleleft P_{j-1})^{\downarrow L}}$$

To define multidimensional distribution it is enough to determine a sequence of oligodimensional distributions

GENERATING SEQUENCE

To define multidimensional distribution it is enough to determine a sequence of oligodimensional distributions

GENERATING SEQUENCE

advantageous are

PERFECT SEQUENCES

they define the same distriibution for both operators

$$\triangleright \text{ and } \triangleleft.$$

Perfect sequences are those, for which

Perfect sequences are those, for which

$$P_1 \triangleright P_2 = P_1 \triangleleft P_2$$

Perfect sequences are those, for which

$$P_1 \triangleright P_2 = P_1 \triangleleft P_2$$

$$P_1 \triangleright P_2 \triangleright P_3 = P_1 \triangleleft P_2 \triangleleft P_3$$

Perfect sequences are those, for which

$$P_1 \triangleright P_2 = P_1 \triangleleft P_2$$

$$P_1 \triangleright P_2 \triangleright P_3 = P_1 \triangleleft P_2 \triangleleft P_3$$

$$P_1 \triangleright P_2 \triangleright P_3 \triangleright P_4 = P_1 \triangleleft P_2 \triangleleft P_3 \triangleleft P_4$$

Perfect sequences are those, for which

$$P_1 \triangleright P_2 = P_1 \triangleleft P_2$$

$$P_1 \triangleright P_2 \triangleright P_3 = P_1 \triangleleft P_2 \triangleleft P_3$$

$$P_1 \triangleright P_2 \triangleright P_3 \triangleright P_4 = P_1 \triangleleft P_2 \triangleleft P_3 \triangleleft P_4$$

$$\vdots$$

$$P_1 \triangleright P_2 \triangleright \ldots \triangleright P_n = P_1 \triangleleft P_2 \triangleleft \ldots \triangleleft P_n$$

Basic properties of perfect sequences

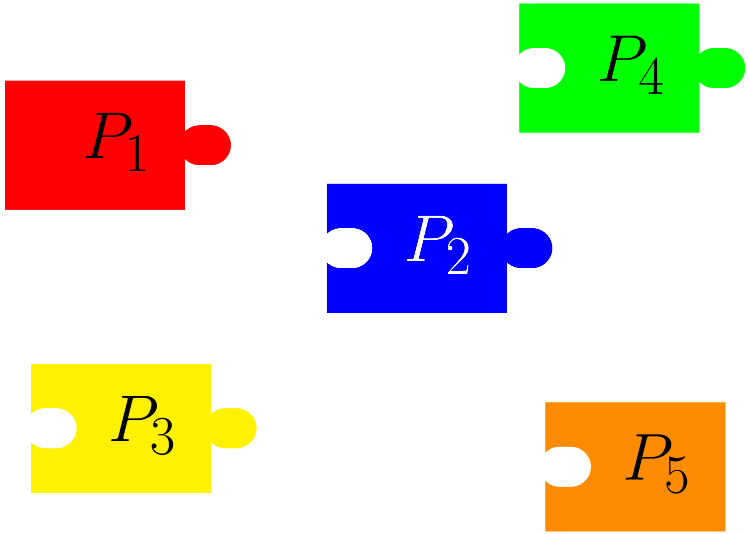
Basic properties of perfect sequences

- $P_1, \ldots, P_m$ is perfect *iff* all the distributions P_j are marginal to $P_1 \triangleright \ldots \triangleright P_m$

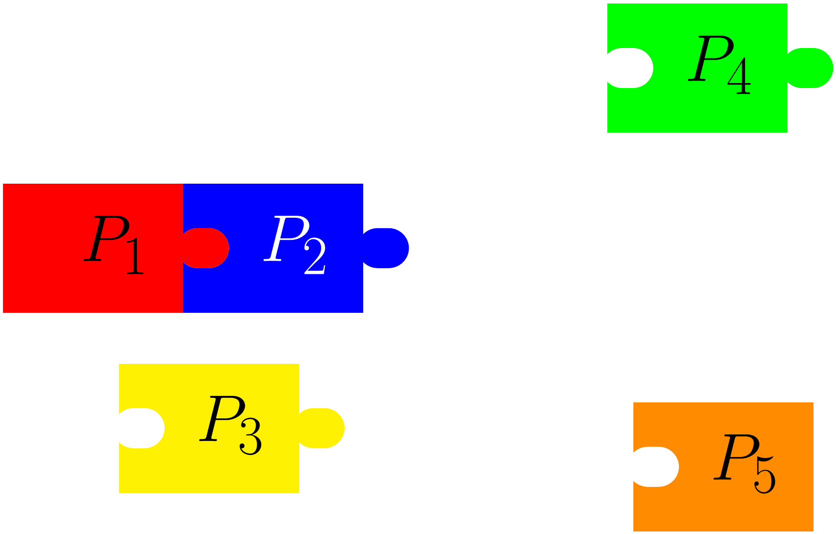
Basic properties of perfect sequences

- $P_1, \ldots, P_m$ is perfect *iff* all the distributions P_j are marginal to $P_1 \triangleright \ldots \triangleright P_m$
- The class Bayesian network models is equivalent to the class of perfect sequence models.

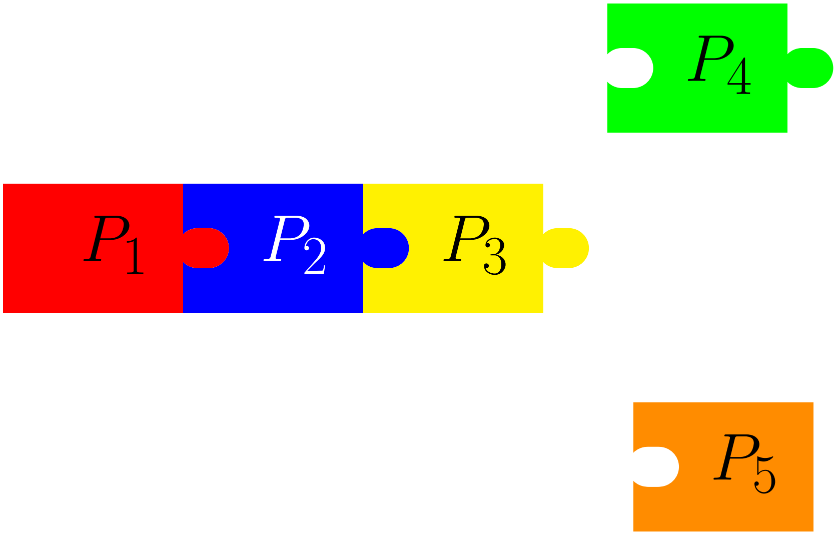
Composition of oligodimensional distributions



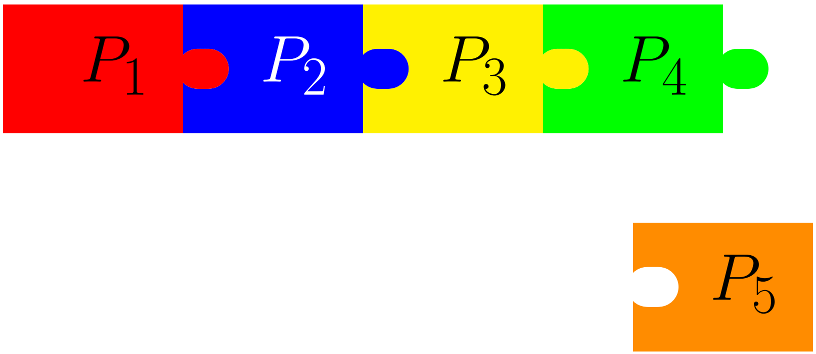
Composition of oligodimensional distributions



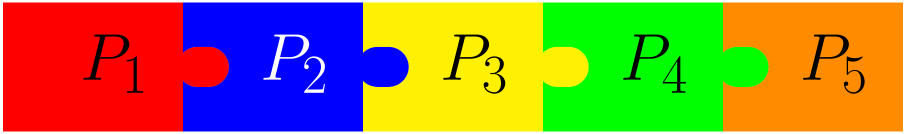
Composition of oligodimensional distributions



Composition of oligodimensional distributions

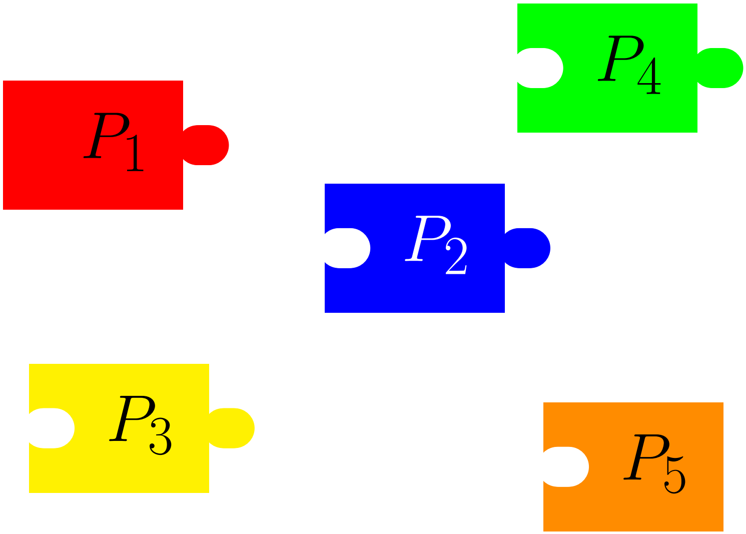


Composition of oligodimensional distributions

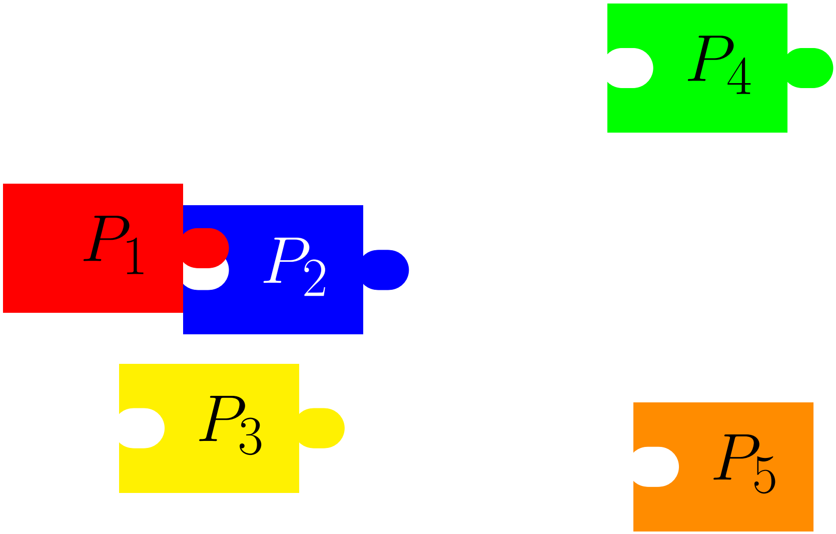


Perfect sequence

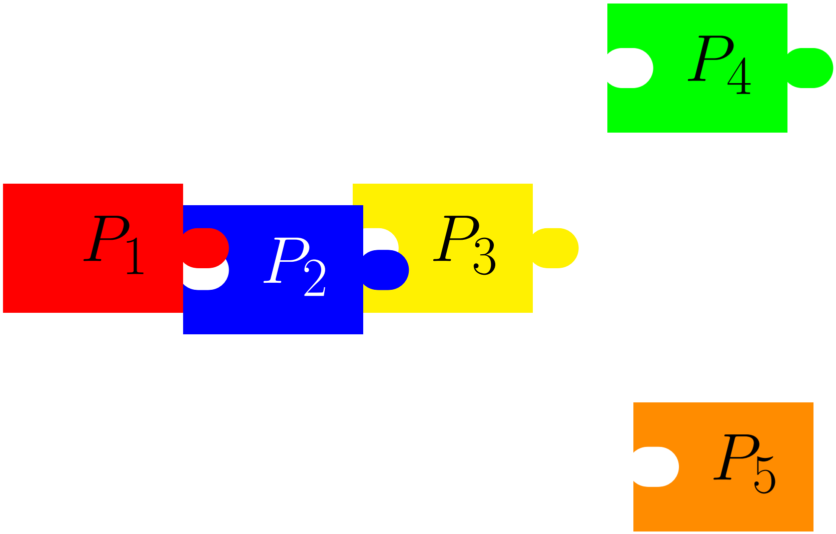
Composition of oligodimensional distributions



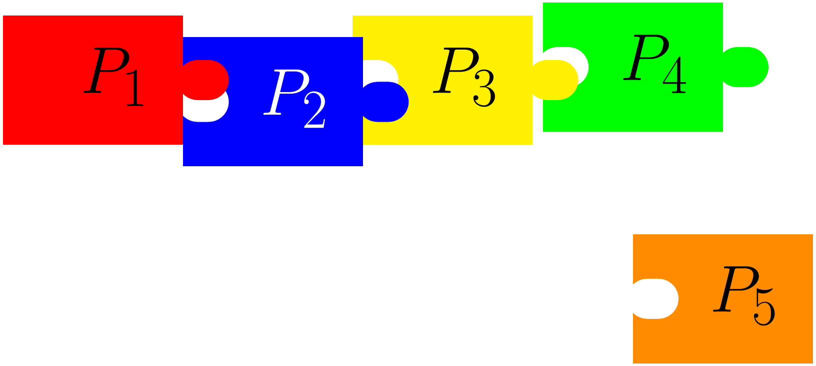
Composition of oligodimensional distributions



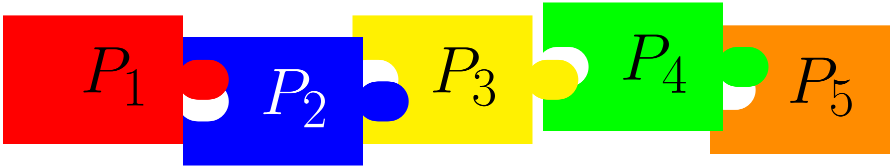
Composition of oligodimensional distributions



Composition of oligodimensional distributions



Composition of oligodimensional distributions



Non-perfect sequence

Comments on equivalence of BN and PS models

- Simple algorithm transforming these types of representation each to other;
- Transformation PS $\longrightarrow$ BN is local;
- Transformation BN $\longrightarrow$ PS is not local (!);
- The difference arises from the fact that in PS models some of marginal distriobutions are explicitely expressed - their computation from BN is sometimes computationally expensive;
- Consequence: simplification of some computational procedures.

What have been achieved in the field of PS models?

- Theoretical apparatus
- Persegrams
- Decomposition
- Efficient marginalization procedure

Comparison of BN and PS models

Shachter's procedure of node deletion and edge reversal

Comparison of BN and PS models

Shachter's procedure of node deletion and edge reversal

- Simple marginalization(!)

Comparison of BN and PS models

Shachter's procedure of node deletion and edge reversal

- Simple marginalization(!)

Bayesian network

terminal node deletion

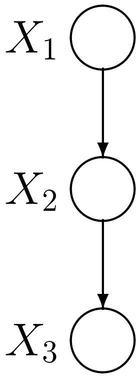
Comparison of BN and PS models

Shachter's procedure of node deletion and edge reversal

- Simple marginalization(!)

Bayesian network

terminal node deletion



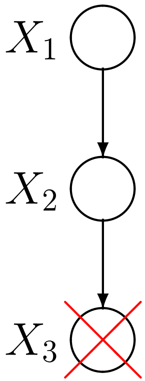
Comparison of BN and PS models

Shachter's procedure of node deletion and edge reversal

- Simple marginalization(!)

Bayesian network

terminal node deletion



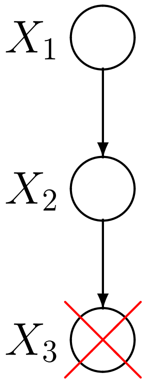
Comparison of BN and PS models

Shachter's procedure of node deletion and edge reversal

- Simple marginalization(!)

Bayesian network | **Compositional model**

terminal node deletion



$$Q_1(X_1, X_2) \triangleright Q_2(X_2, X_3)$$

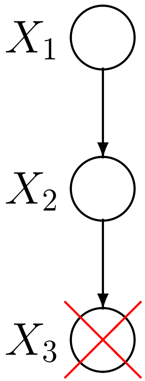
Comparison of BN and PS models

Shachter's procedure of node deletion and edge reversal

- Simple marginalization(!)

Bayesian network

terminal node deletion



Compositional model

deletion of a variable appearing among the arguments of only one distribution

$$Q_1(X_1, X_2) \triangleright Q_2(X_2, X_3)$$

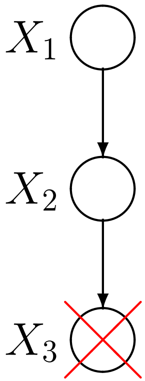
Comparison of BN and PS models

Shachter's procedure of node deletion and edge reversal

- Simple marginalization(!)

Bayesian network

terminal node deletion



Compositional model

deletion of a variable appearing among the arguments of only one distribution

$$Q_1(\cancel{X_1}, X_2) \triangleright Q_2(X_2, \cancel{X_3})$$

Decomposition of PS models

Decomposition of PS models

$$P_1(x_1, x_2), P_2(x_1, x_3), P_3(x_3, x_4, x_7), P_4(x_2, x_5), P_5(x_3, x_6), P_6(x_5, x_6, x_9),$$
$$P_7(x_6, x_7, x_{11}), P_8(x_5, x_8), P_9(x_6, x_{10}), P_{10}(x_8, x_9, x_{10}, x_{12})$$

Decomposition of PS models

$$P_1(x_1, x_2), P_2(x_1, x_3), P_3(x_3, x_4, x_7), P_4(x_2, x_5), P_5(x_3, x_6), P_6(x_5, x_6, x_9),$$
$$P_7(x_6, x_7, x_{11}), P_8(x_5, x_8), P_9(x_6, x_{10}), P_{10}(x_8, x_9, x_{10}, x_{12})$$

$P_1, P_2, P_4, P_5, P_6, P_8, P_9, P_{10}$

P_3, P_5, P_7

Decomposition of PS models

$$P_1(x_1, x_2), P_2(x_1, x_3), P_3(x_3, x_4, x_7), P_4(x_2, x_5), P_5(x_3, x_6), P_6(x_5, x_6, x_9),$$
$$P_7(x_6, x_7, x_{11}), P_8(x_5, x_8), P_9(x_6, x_{10}), P_{10}(x_8, x_9, x_{10}, x_{12})$$

$P_1, P_2, P_4, P_5, P_6, P_8, P_9, P_{10}$

P_3, P_5, P_7

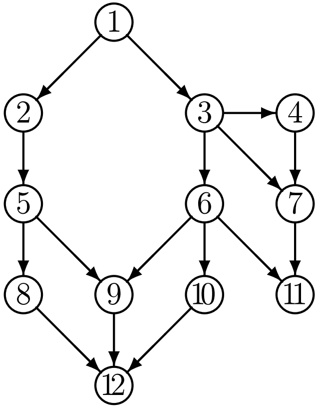
$P_1, P_2, P_4, P_5, P_6,$

P_6, P_8, P_9, P_{10}

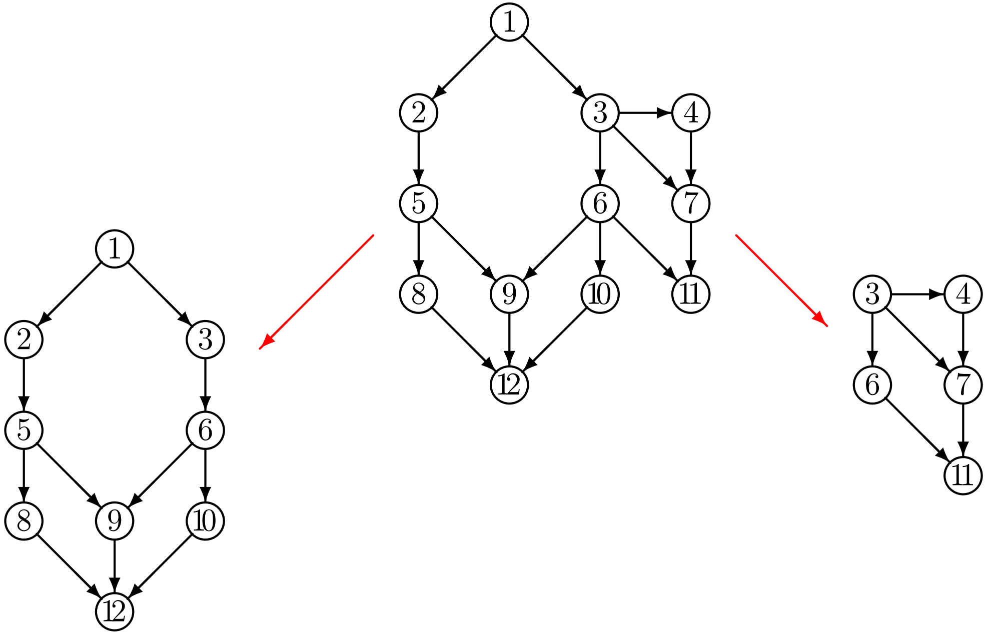
$P_3,$

$P_3^{\{(3,7)\}}, P_5, P_7$

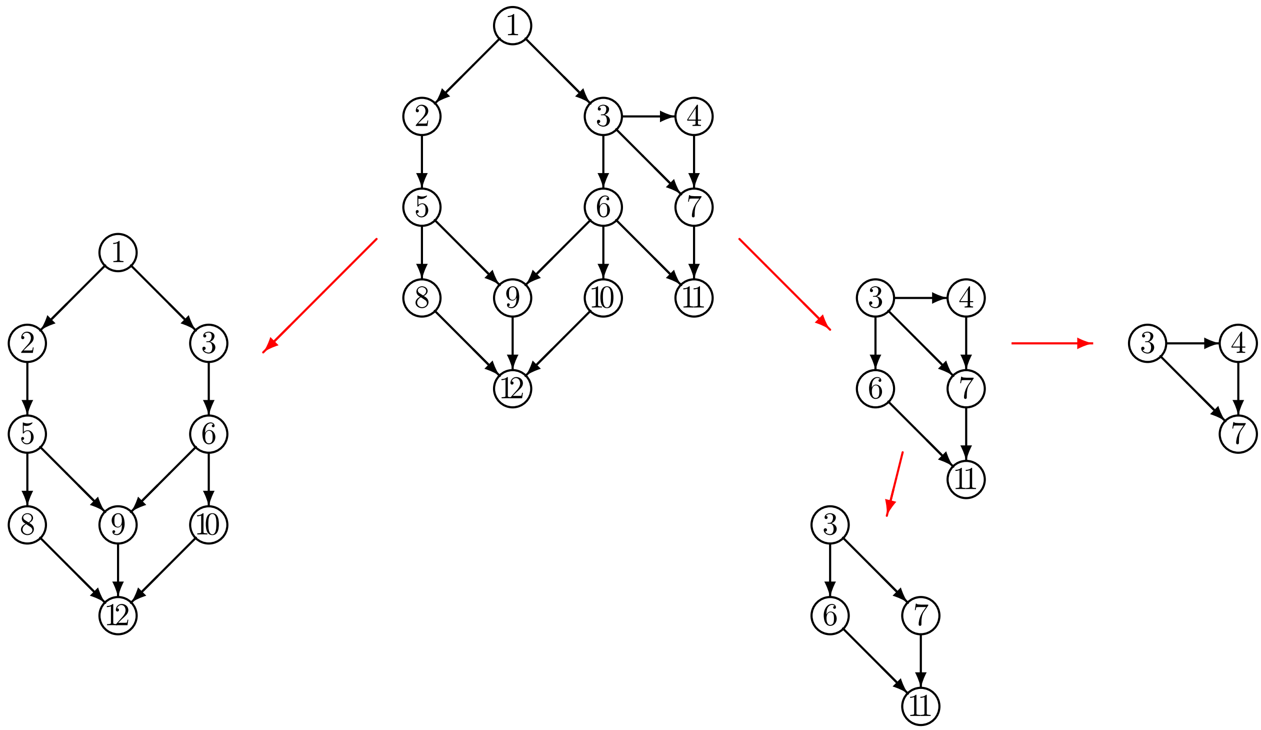
Decomposition of Bayesian networks



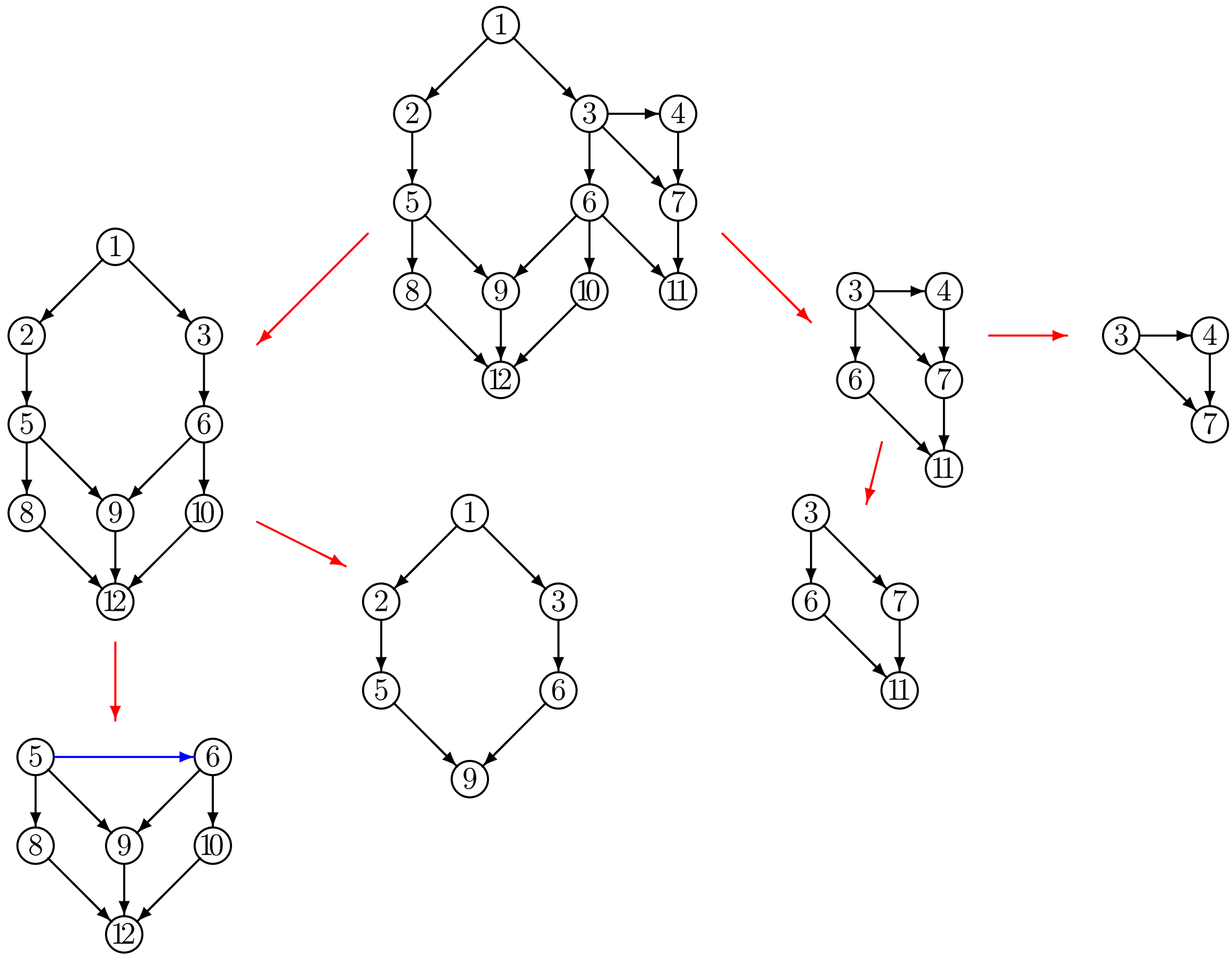
Decomposition of Bayesian networks



Decomposition of Bayesian networks



Decomposition of Bayesian networks



Concluding remark

Concluding remark

Based on Albert Perez' ideas of dependence structure siplification approximation we have developed an approach for multidimensional distributions modelling, which is an alternative to Bayesian networks and which is from some points of view superioir to them.