Decomposition

Basic Idea

- A difficult problem is broken down into small sub-problems, which can then be solved individually and combined to form an overall result.

- For example, if you want to write a book, you can write a sketch as a framework, then approach each component individually and finally put everything together to form a coherent work.

- Similar to the “divide and conquer” method for algorithms.

Real World Example

Property family	Car body	Motor	Radio	Doors	Seat cover	Makeup mirrow	…
Property	Hatch-back	2.8 L 150 kW Otto	Type alpha	4	Leather, Type L3	yes	…

About 200 variables

Typically 4 to 8, but up to 150 instances per variable

More than 2^{200} combinations exist,

but lots of combinations are not **possible.**

Real World Example : Planning in car manufacturing

Available information: 10000 technical rules, 200 attributes

"If Motor = m_4 and Heating = h_1 then Generator $\in \{g_1, g_3, g_5\}$"

"Engine type e_1 can only be combined with transmission t_2 or t_5."

"Transmission t_5 requires crankshaft c_2."

"Convertibles have the same set of radio options as SUVs."

Each information corresponds to a constraint in a high dimensional subspace, possible questions/inferences:

"Can a station wagon with engine e_4 be equipped with tire set y_6?"

"Supplier S_8 failed to deliver on time. What production line has to be modified and how?"

"Are there any peculiarities within the set of cars that suffered an aircondition failure?"

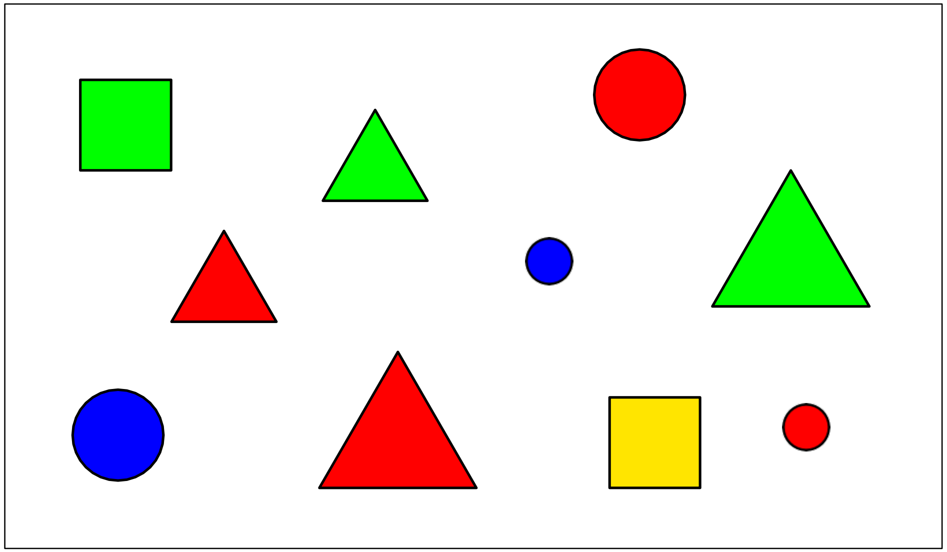
Handling a Problem by Decomposition

Given: A large (high-dimensional) δ representing the domain knowledge.

Desired: A set of smaller (lower-dimensional) $\{\delta_1, \ldots, \delta_s\}$ (maybe overlapping) from which the original δ *could* be reconstructed with no (or as few as possible) errors.

With such a decomposition we can draw any conclusions from $\{\delta_1, \ldots, \delta_s\}$ that could be inferred from δ— without, however, actually reconstructing it.

Toy World



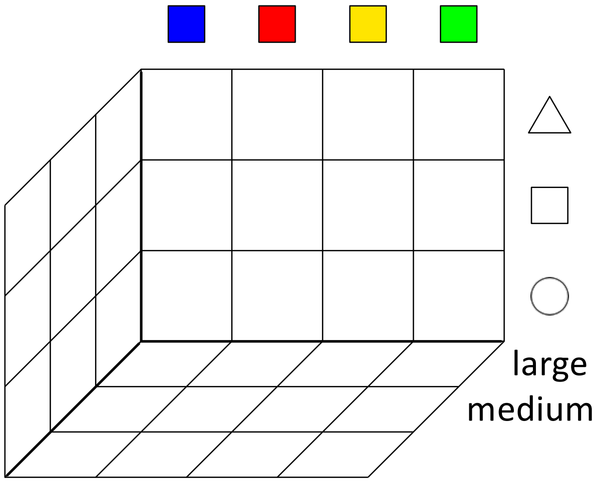
Relation

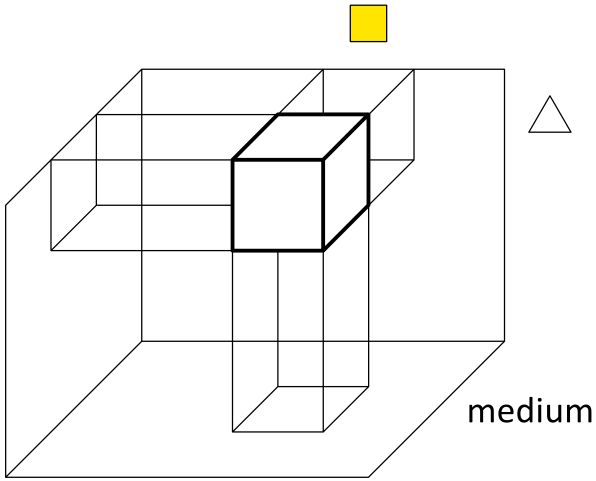
color	shape	size
🟦	○	small
🟦	○	medium
🟥	○	small
🟥	○	medium
🟥	△	medium
🟥	△	large
🟨	□	medium
🟩	□	medium
🟩	△	medium
🟩	△	large

- 10 simple geometric objects, 3 attributes
- One object is chosen at random and examined
- Inferences are drawn about the unobserved attributes

Note: In real applications the attributes (variables) could be motor, heating, generator, etc.

Example 1: The Reasoning Space (Frame of Discernment)





The reasoning space consists of a finite set E of states.

The states are described by a set of n attributes A_i, $i = 1, \ldots, n$, whose domains $\{a_1^{(i)}, \ldots, a_{n_i}^{(i)}\}$ can be seen as sets of propositions or events.

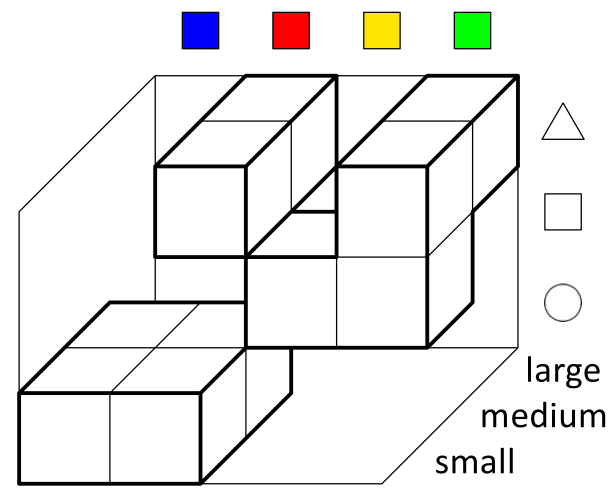
The events in a domain are mutually exclusive and exhaustive.

Example 1:The Relation in the Reasoning Space

Relation

color	shape	size
🟦	○	small
🟦	○	medium
🟥	○	small
🟥	○	medium
🟥	△	medium
🟥	△	large
🟨	□	medium
🟩	□	medium
🟩	△	medium
🟩	△	large

Visual Description



Each cube represents one tuple.

The spatial representation helps to understand the decomposition mechanism.

Possibility-Based Formalization of Reasoning Space

Definition: Let E be a (finite) sample space.
A discrete possibility measure R on E is a function $R : 2^{E} \rightarrow \{0,\ 1\}$ satisfying

1. $R(\emptyset) = 0$ and
2. $\forall A_1, A_2 \subseteq E : R(A_1 \cup A_2) = \max\{R(A_1),\ R(A_2)\}$.

Similar to Kolmogorov's axioms of probability theory.

If an event A can occur (if it is possible), then $R(A) = 1$,
otherwise (if A cannot occur/is impossible) then $R(A) = 0$.

$R(E) = 1$ is not required, because this would exclude the empty relation.

From the axioms it follows $R(E_1 \cap E_2) \leq \min\{R(E_1),\ R(E_2)\}$.

Note: In our course, only the possibility degrees 0 and 1 are used. A general possibility measure can have value in the unit interval.

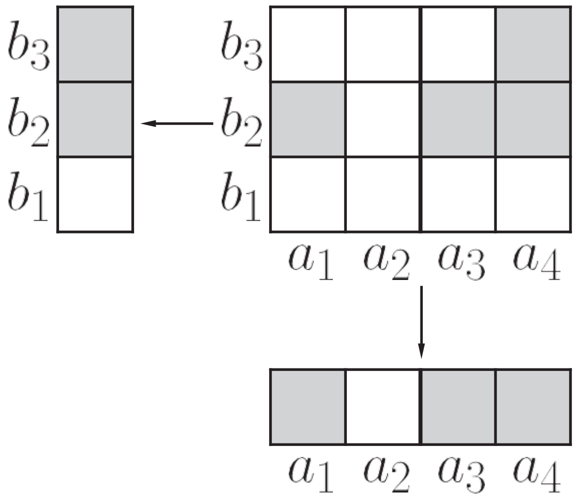
Projection / Marginalization

Let R_{AB} be a relation over two attributes A and B. The projection (or marginalization) from schema $\{A, B\}$ to schema $\{A\}$ is defined as:

$$\forall a \in \mathrm{dom}(A): \quad R_A(A = a) = \max_{b \in \mathrm{dom}(B)} \{R_{AB}(A = a, B = b)\}$$

Note: dom(B)=domain (B) =range (B), set of possible values of the variable B.

This principle is easily generalized to sets of attributes.

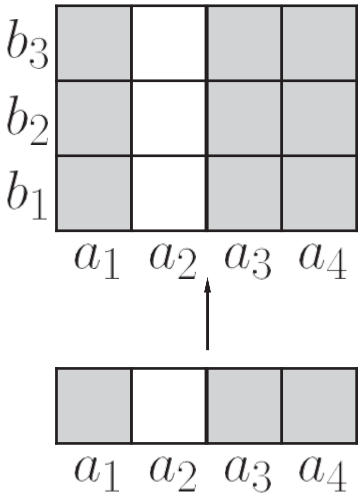


Cylindrical Extension

Let R_A be a relation over an attribute A. The cylindrical extension R_{AB} from $\{A\}$ to $\{A, B\}$ is defined as:

$$\forall a \in \mathrm{dom}(A) : \forall b \in \mathrm{dom}(B) : \quad R_{AB}(A = a, B = b) \quad = \quad R_A(A = a)$$

This principle is easily generalized to sets of attributes.



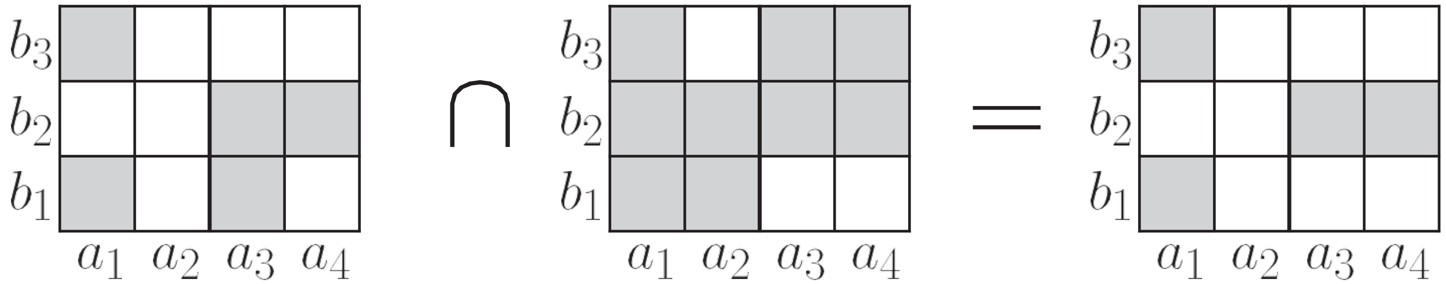
Intersection

Let $R^{(1)}_{AB}$ and $R^{(2)}_{AB}$ be two relations with attribute schema $\{A, B\}$. The intersection

R_{AB} of both is defined in the natural way:

$$\forall a \in \mathrm{dom}(A) : \forall b \in \mathrm{dom}(B) :$$
$$R_{AB}(A = a, B = b) = \min\{R^{(1)}_{AB}(A = a, B = b), R^{(2)}_{AB}(A = a, B = b)\}$$

This principle is easily generalized to sets of attributes.

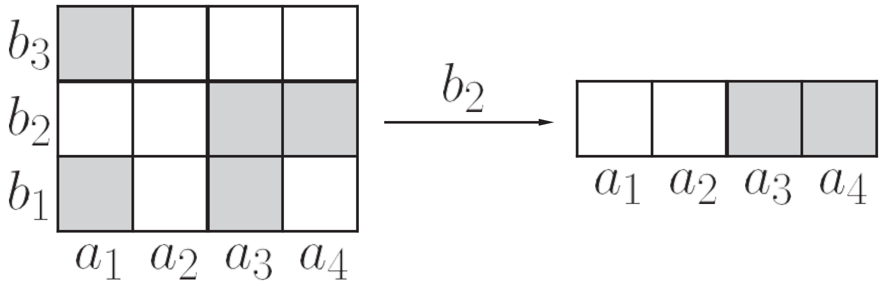


Conditional Relation

Let R_{AB} be a relation over the attribute schema $\{A,\ B\}$. The conditional relation of A given B is defined as follows:

$$\forall a \in \mathrm{dom}(A) : \forall b \in \mathrm{dom}(B) : R_A(A = a \mid B = b) = R_{AB}(A = a, B = b)$$

This principle is easily generalized to sets of attributes.

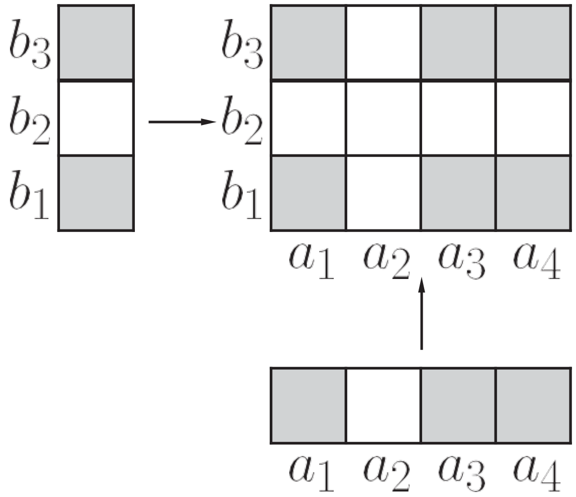


(Unconditional) Independence

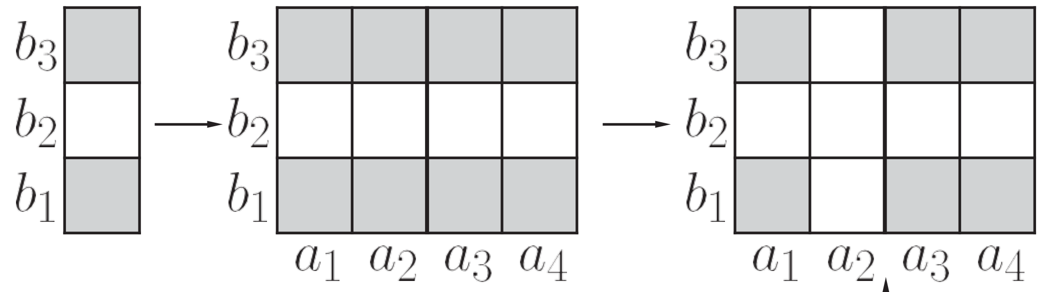
Let R_{AB} be a relation over the attribute schema $\{A, B\}$. We call A and B relationally independent (w. r. t. R_{AB}) if the following condition holds:

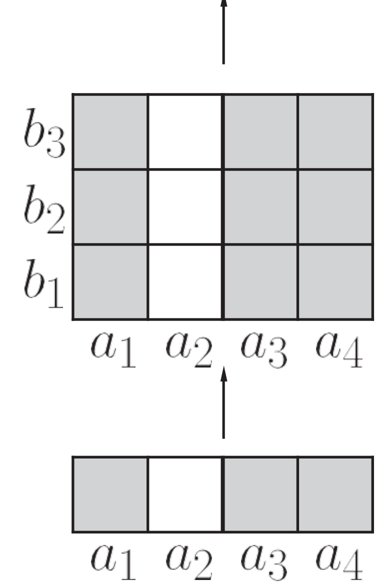
$$\forall a \in \text{dom}(A) : \forall b \in \text{dom}(B) : R_{AB}(A = a, B = b) = \min\{R_A(A = a), R_B(B = b)\}$$

This principle is easily generalized to sets of attributes.



(Unconditional) Independence





Intuition: Fixing one (possible) value of A does not restrict the (possible) values of B and vice versa.

Conditioning on any possible value of B always results in the same relation R_A.

Alternative independence expression:

$$\forall b \in \mathrm{dom}(B): \quad R_B(B = b) = 1: \quad R_A(A = a \mid B = b) \quad = \quad R_A(A = a)$$

Decomposition

The original two-dimensional relation can be reconstructed from the two one-dimensional ones, if we have (unconditional) independence.

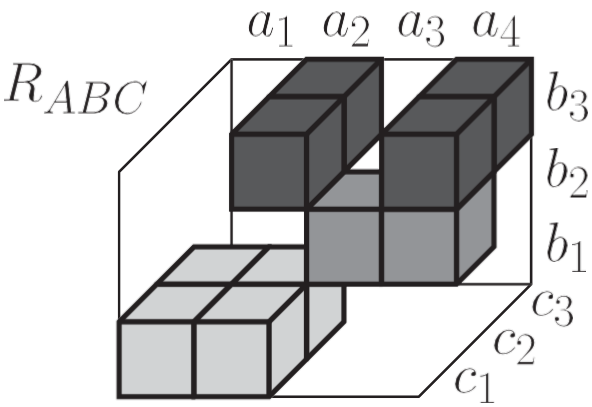
The definition for (unconditional) independence already told us how to do so:

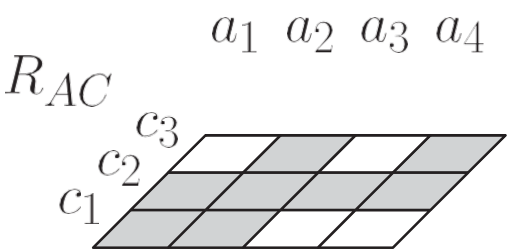
$$R_{AB}(A = a, B = b) \quad = \quad \min\{R_A(A = a), R_B(B = b)\}$$

Storing R_A and R_B is sufficient to represent the information of R_{AB}.

Question: The (unconditional) independence is a rather strong restriction. Are there other types of independence that allow for a decomposition as well?

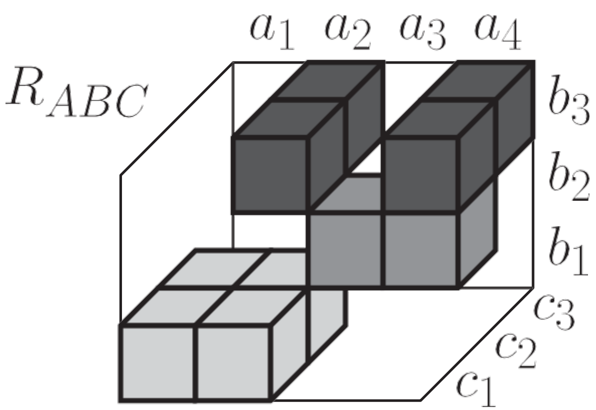
Conditional Relational Independence

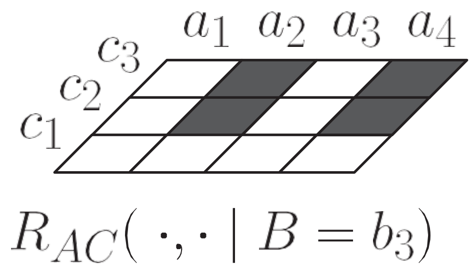


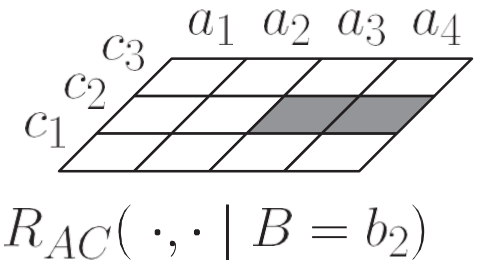


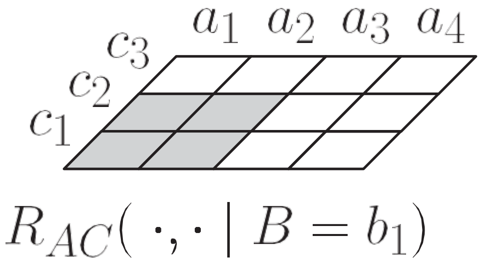
Clearly, A and C are unconditionally dependent, i. e. the relation R_{AC} cannot be reconstructed from R_A and R_C.

Conditional Relational Independence









However, given all possible values of B, all respective conditional relations R_{AC} show the independence of A and C.

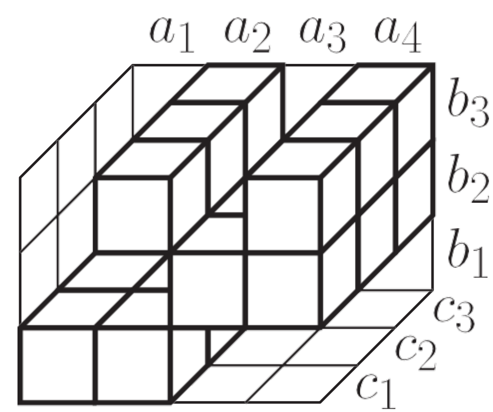
$$R_{AC}(a,c \mid b) = \min\{R_A(a \mid b), R_C(c \mid b)\}$$

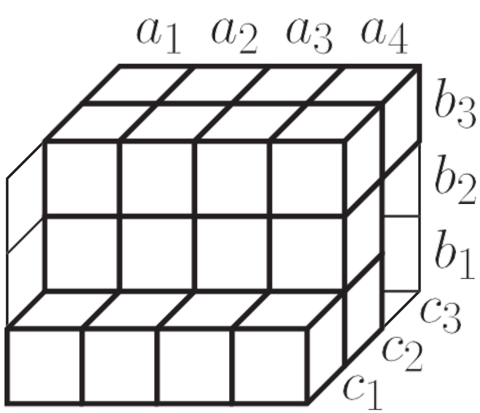
With the definition of a conditional relation, the de-composition description for R_{ABC} reads:

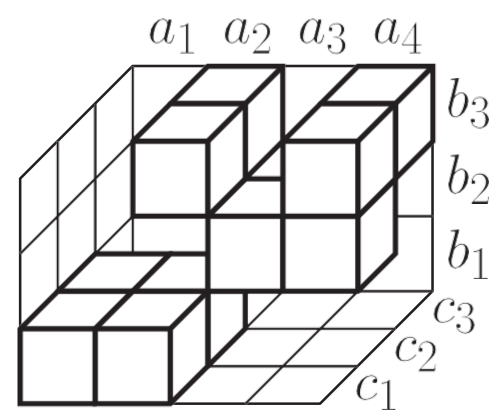
$$R_{ABC}(a,b,c) = \min\{R_{AB}(a,b), R_{BC}(b,c)\}$$

Again, we reconstruct the initial relation from the cylindrical extentions of the two relations formed by the attributes A, B and B, C.

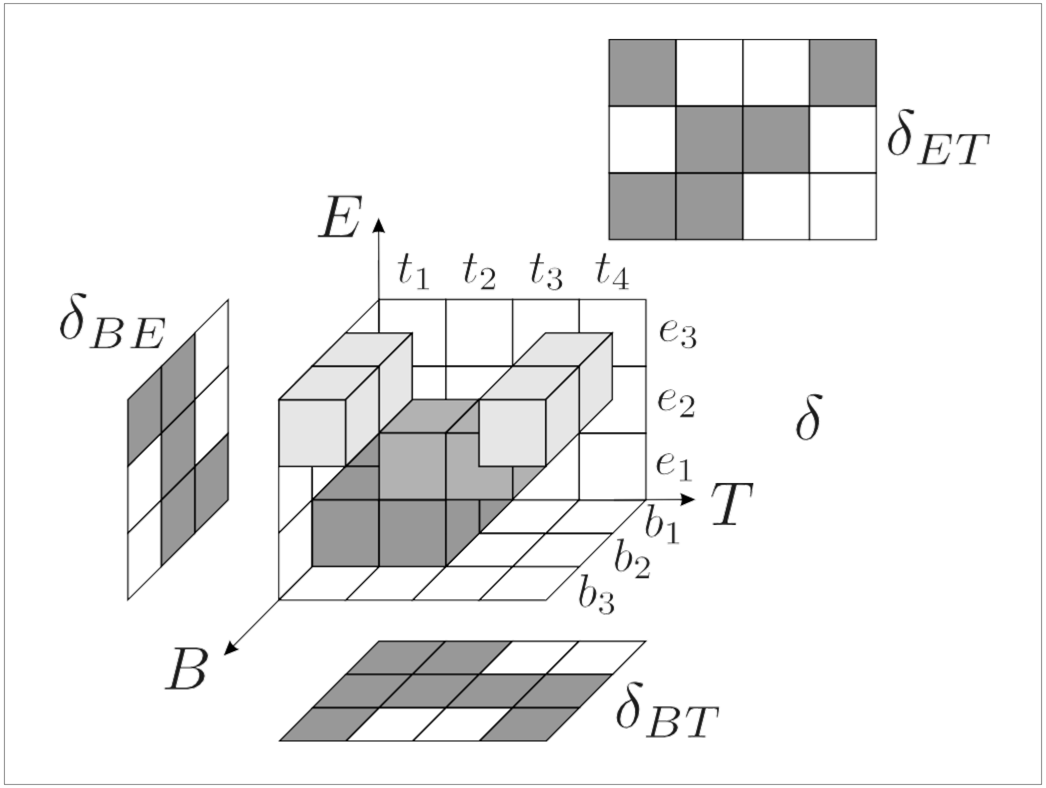
It is possible since A and C are (relationally) conditionally independent given B.





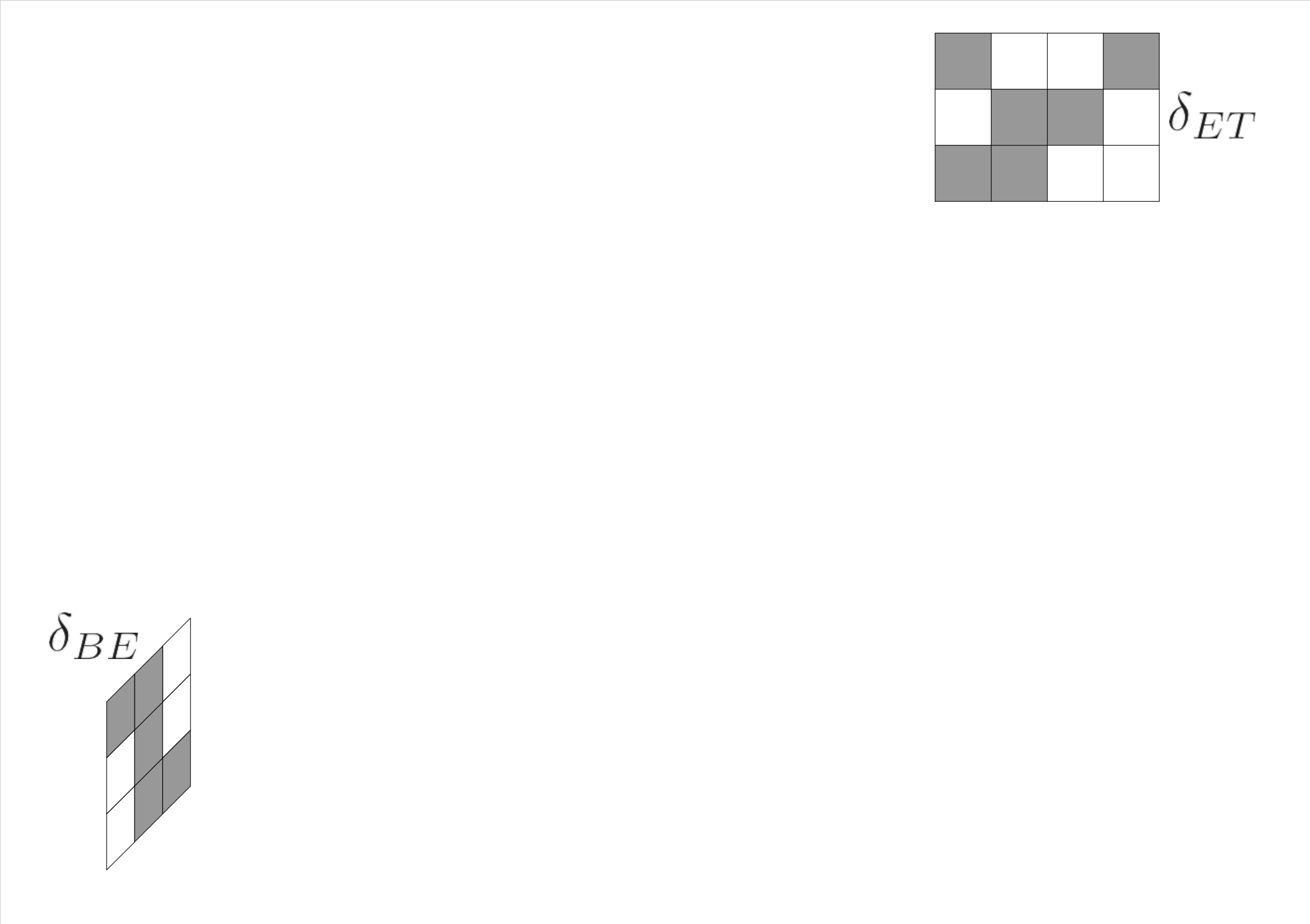


Example 2: Projections

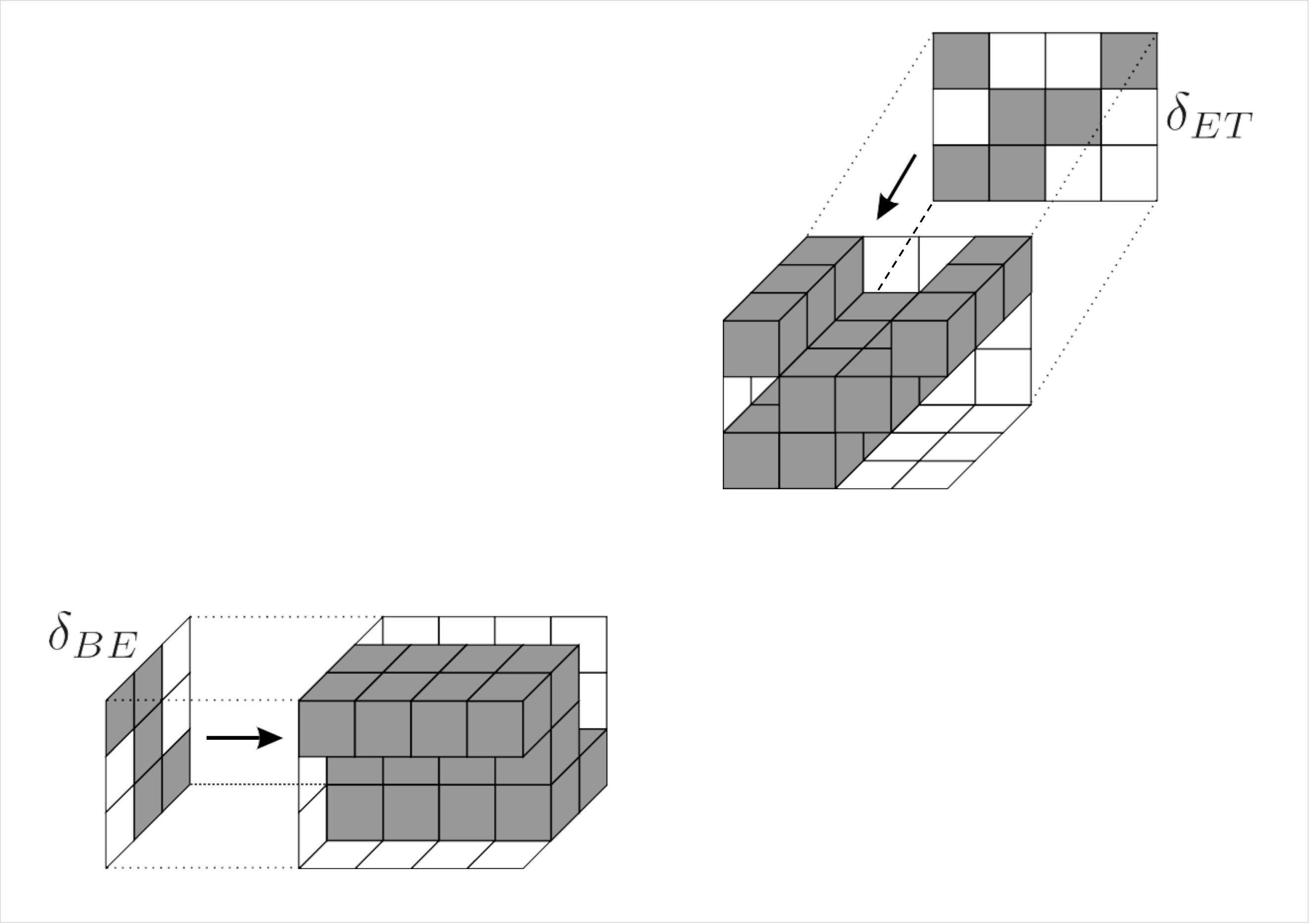


Is it possible to reconstruct δ from the three projections?

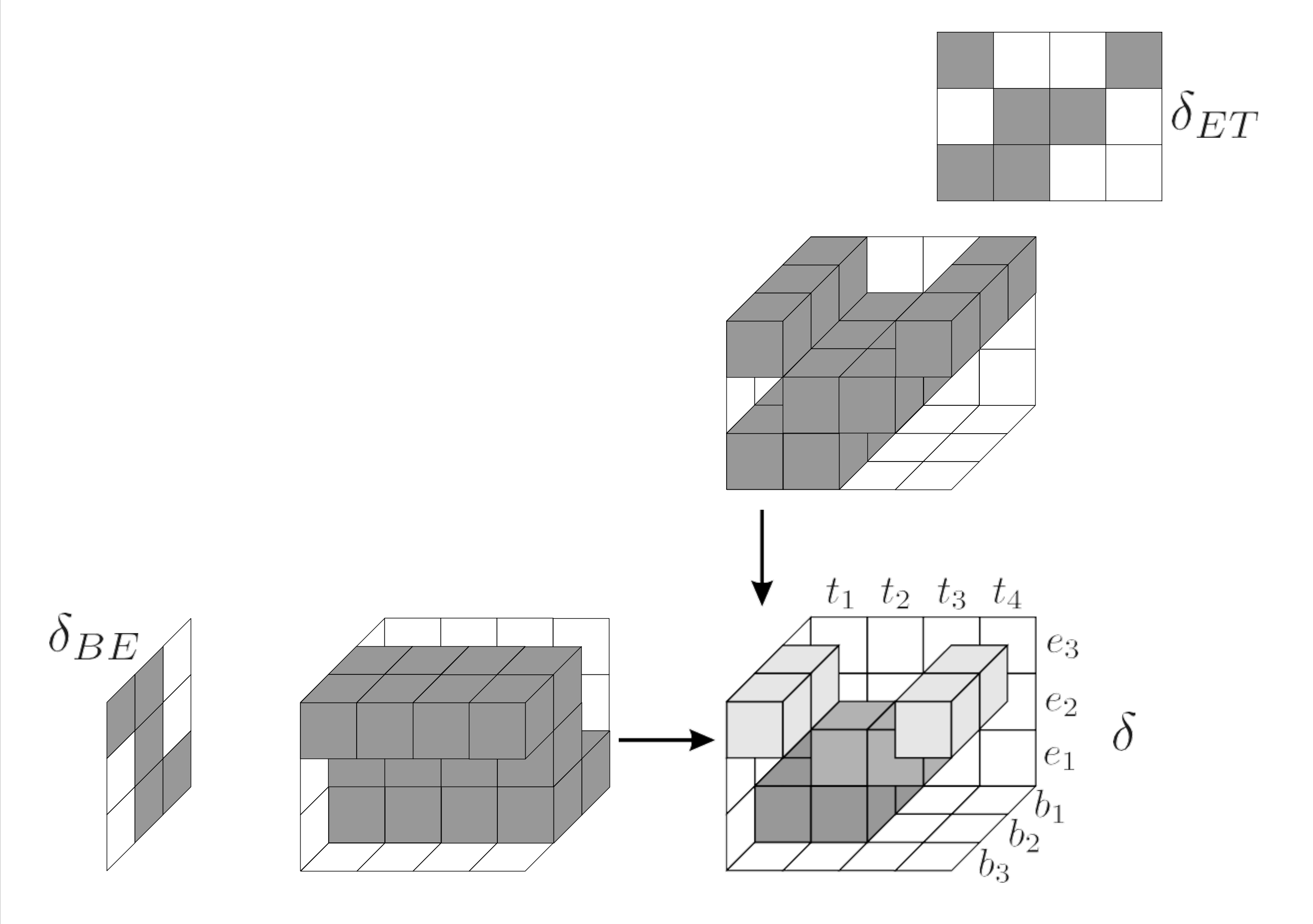
Example2: Reconstruction of δ with δ_{BE} and δ_{ET}



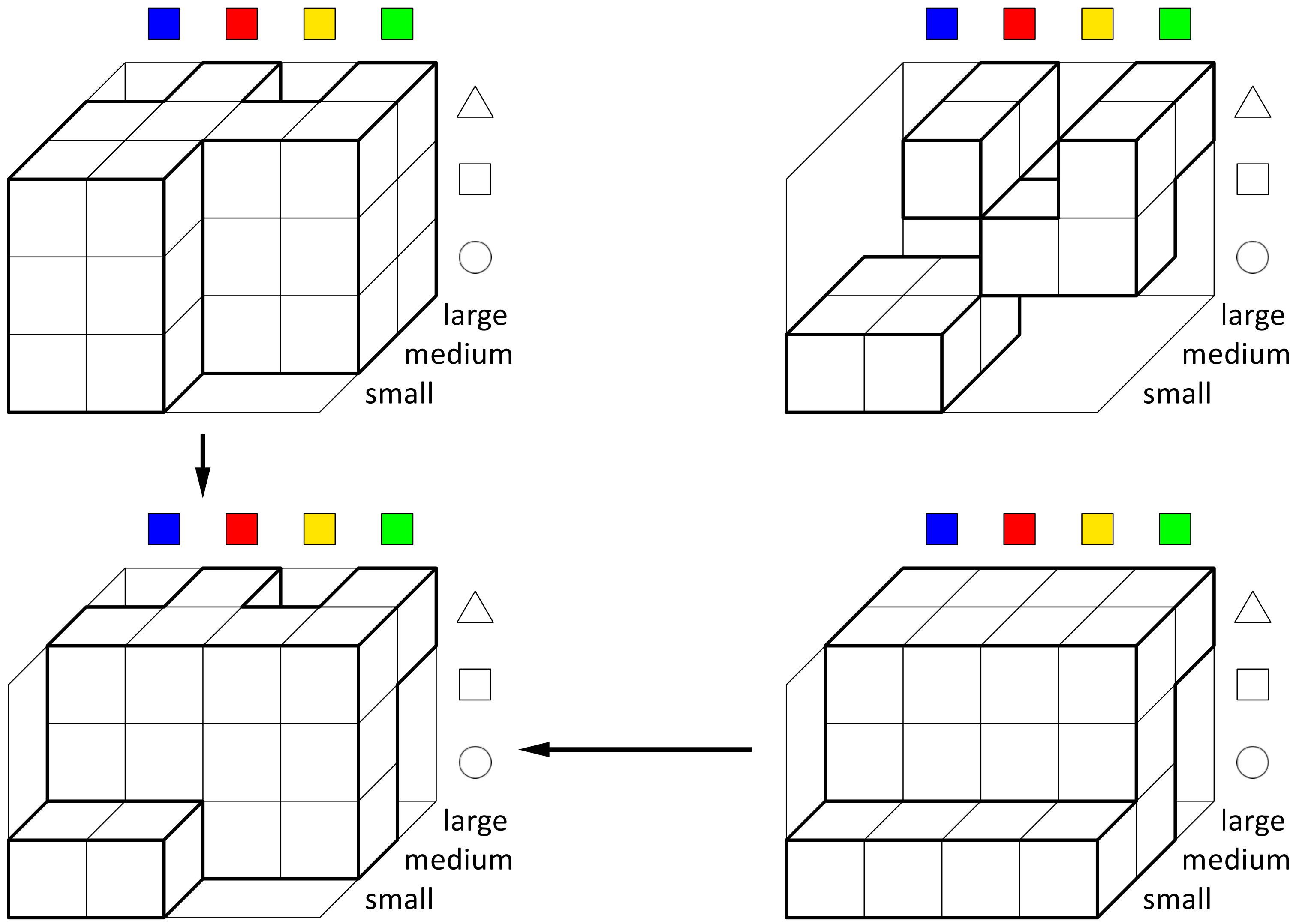
Example2:Reconstruction of δ with δ_{BE} and δ_{ET}



Example 2: Reconstruction of δ with δ_{BE} and δ_{ET}

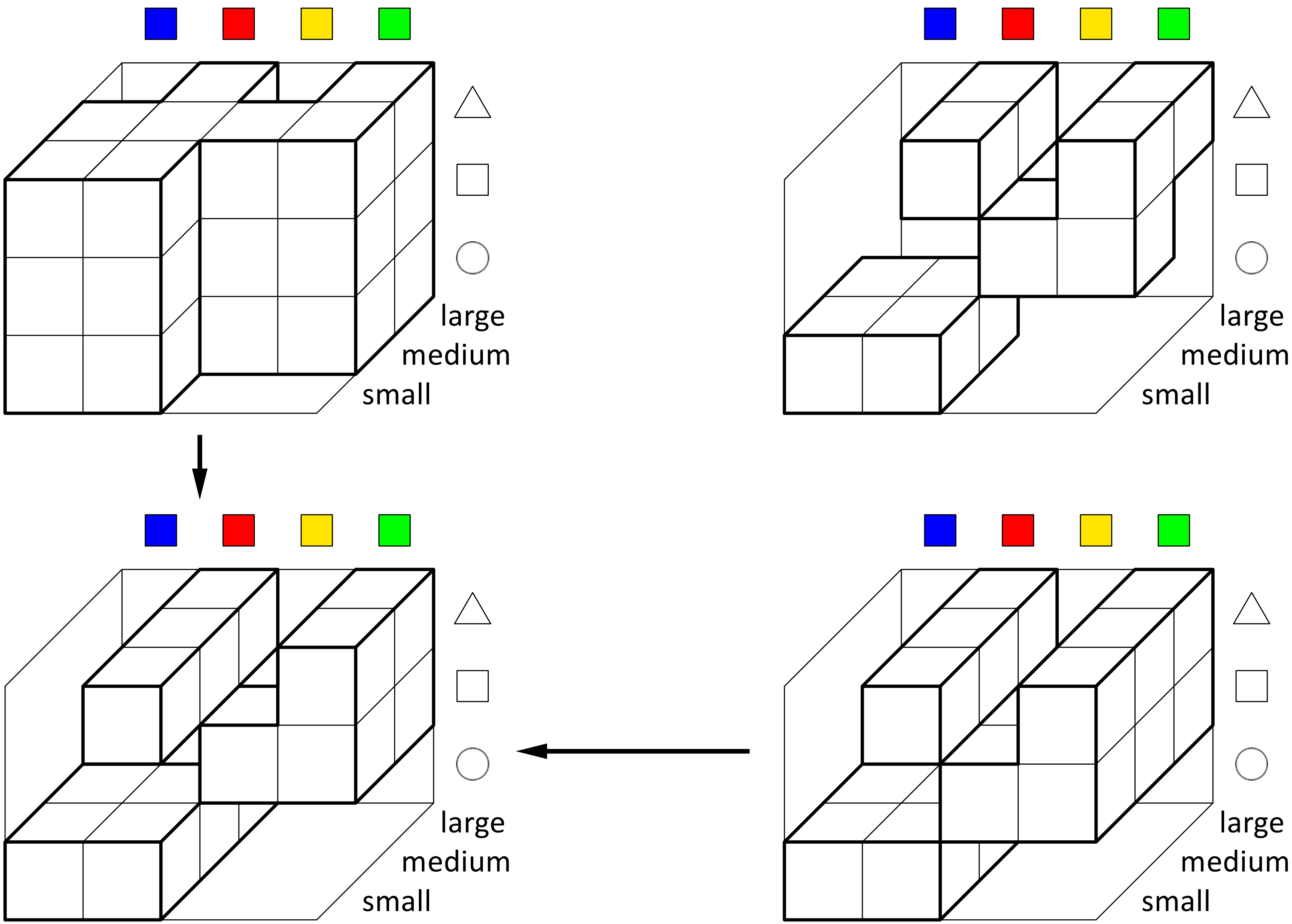


Example 3: Using other Projections 1



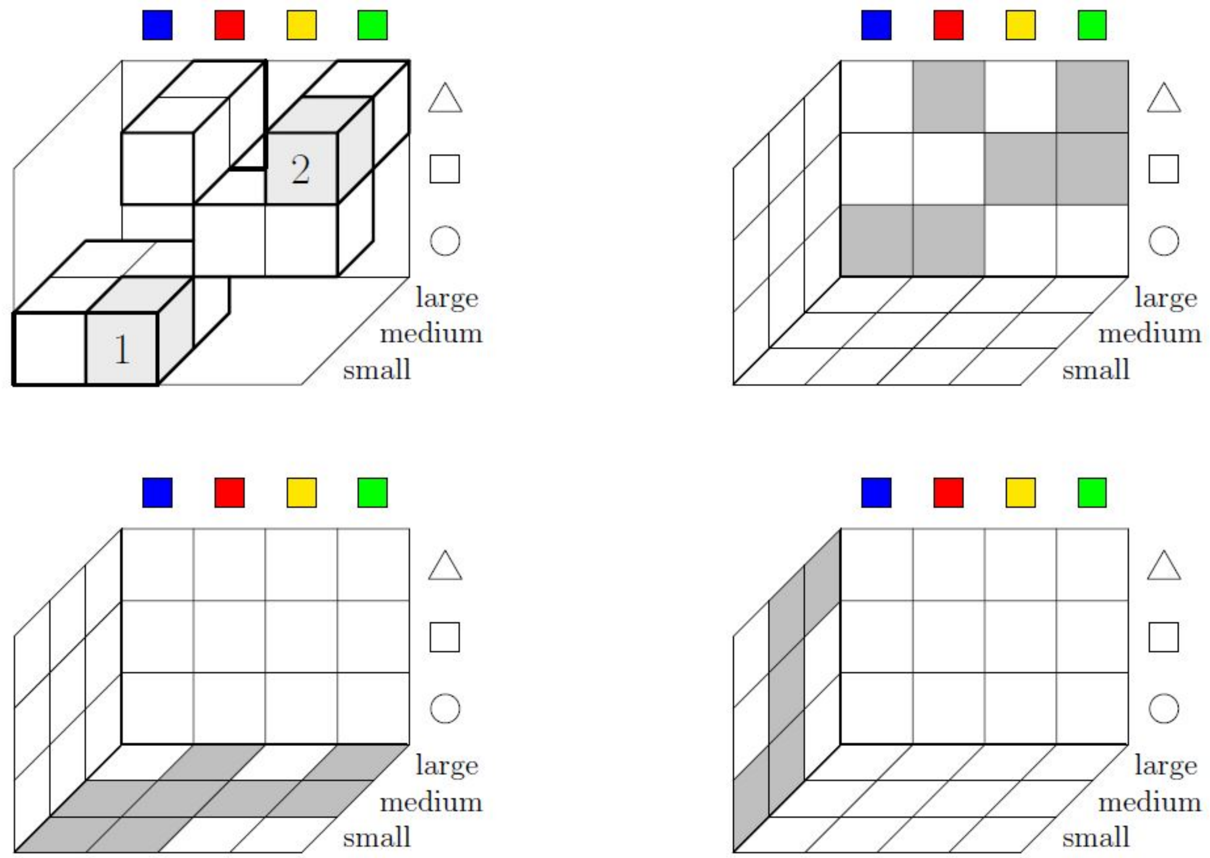
This choice of subspaces does not yield a decomposition.

Example 3: Using other Projections 2



This choice of subspaces does not yield a decomposition.

Example 3: Is Decomposition Always Possible?



A modified relation (without tuples 1 or 2) may not possess a decomposition.

Possibility-Based Formalization of Decomposition

Definition: Let $U = \{A_1, \ldots, A_n\}$ be a set of attributes and r_U a relation over U. Furthermore, let $\mathcal{M} = \{M_1, \ldots, M_m\} \subseteq 2^U$ be a set of nonempty (but not necessarily disjoint) subsets of U satisfying

$$\bigcup_{M \in \mathcal{M}} M = U.$$

r_U is called **decomposable** w.r.t. $\mathcal{M}$ iff

$$\forall a_1 \in \mathrm{dom}(A_1) : \ldots \forall a_n \in \mathrm{dom}(A_n) :$$
$$r_U\left(\bigwedge_{A_i \in U} A_i = a_i\right) = \min_{M \in \mathcal{M}} \left\{ r_M\left(\bigwedge_{A_i \in M} A_i = a_i\right)\right\}.$$

If r_U is decomposable w.r.t. $\mathcal{M}$, the set of relations

$$\mathcal{R}_{\mathcal{M}} = \{r_{M_1}, \ldots, r_{M_m}\} = \{r_M \mid M \in \mathcal{M}\}$$

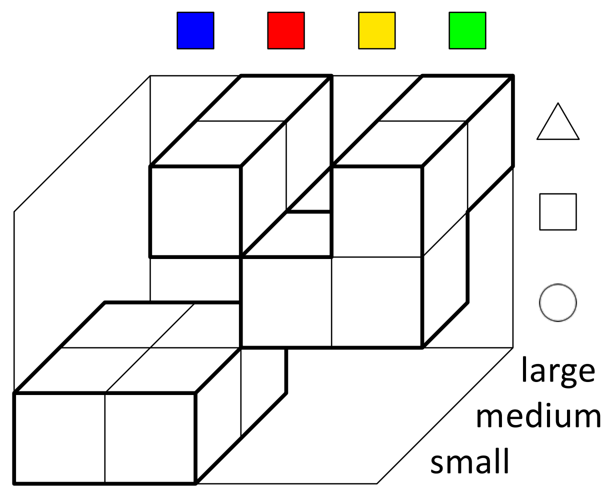
is called the **decomposition** of r_U.

Equivalent to **join decomposability** in database theory (natural join).

Relation

color	shape	size
🟦	○	small
🟦	○	medium
🟥	○	small
🟥	○	medium
🟥	△	medium
🟥	△	large
🟨	□	medium
🟩	□	medium
🟩	△	medium
🟩	△	large

Spatial Visualisation



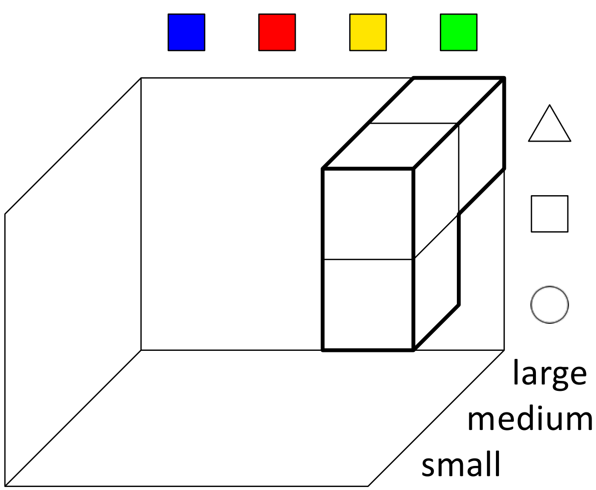
Each cube represents one tuple.

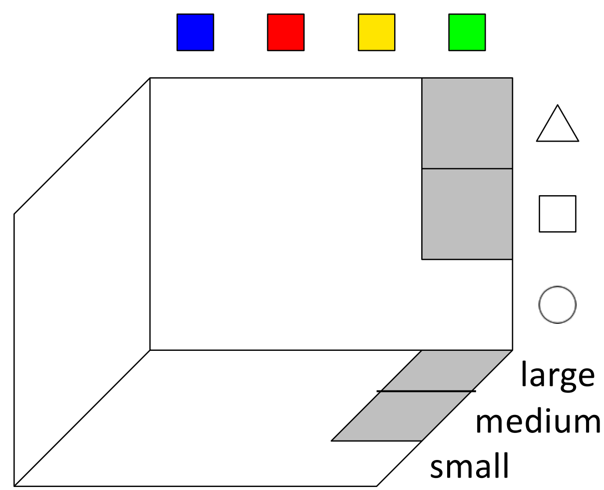
The spatial representation helps to understand the decomposition mechanism.

Let it be known (e.g. from an observation) that the given object is green.

This observation considerably reduces the space of possible value combination: It follows that the given object must be

- either a triangle or a square and
- either medium or large.

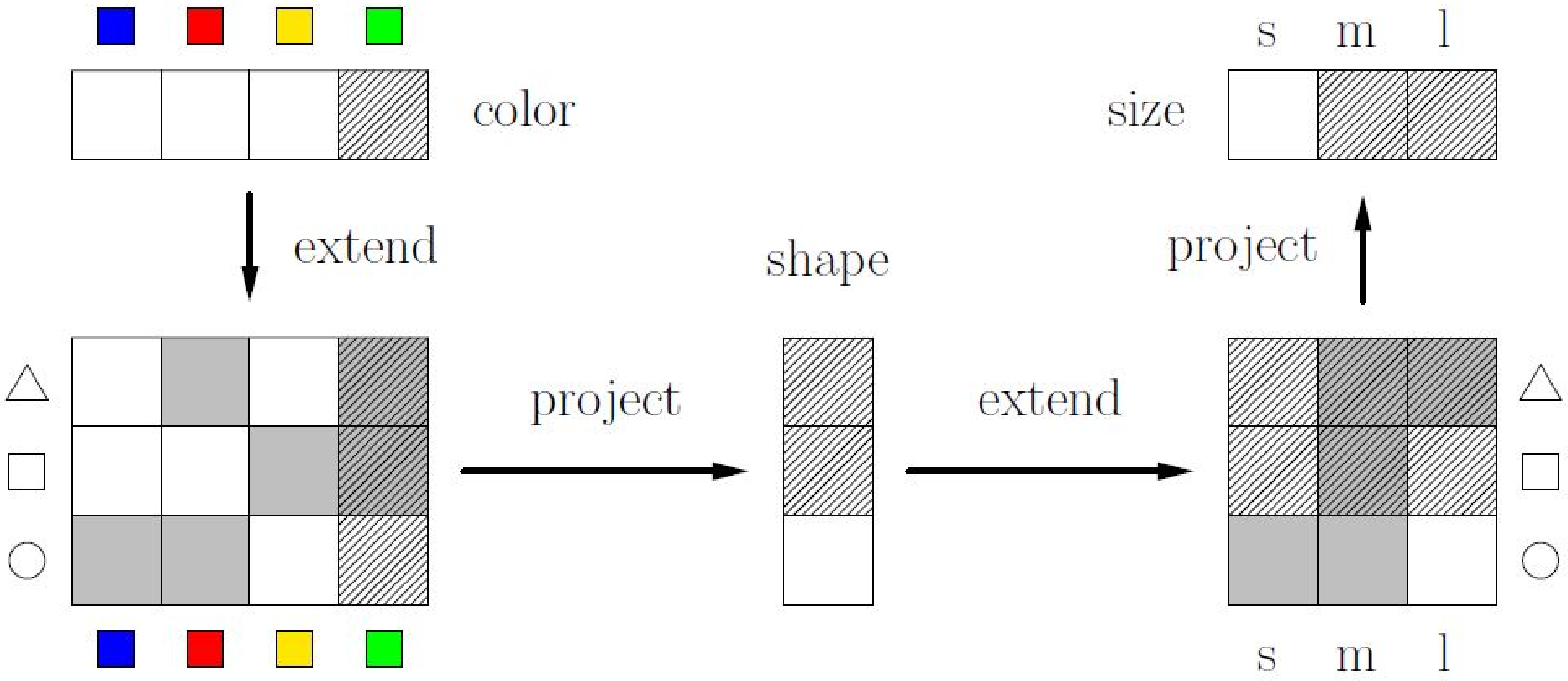




Note that (formulated in the language of Data Science), evidence was used for updating of our a priori knowledge. We can use now the more informative, so called a posteriori knowledge.

Example 4: Relational Evidence Propagation

Due to the fact that color and size are conditionally independent given the shape, the reasoning result can be obtained using only the projections to the subspaces:



This reasoning scheme can be formally justified with discrete possibility measures.

Example 4: Relational Evidence Propagation, Step 1

$$
\begin{aligned}
& R(B = b \mid A = a_{\text{obs}}) \\
= \; & R\left(\bigvee_{a \in \text{dom}(A)} A = a, B = b, \bigvee_{c \in \text{dom}(C)} C = c \,\middle|\, A = a_{\text{obs}} \right) \\
\overset{(1)}{=} \; & \max_{a \in \text{dom}(A)} \Big\{ \max_{c \in \text{dom}(C)} \{ R(A = a, B = b, C = c \mid A = a_{\text{obs}}) \} \Big\} \\
\overset{(2)}{=} \; & \max_{a \in \text{dom}(A)} \Big\{ \max_{c \in \text{dom}(C)} \{ \min\{ R(A = a, B = b, C = c), R(A = a \mid A = a_{\text{obs}}) \} \} \Big\} \\
\overset{(3)}{=} \; & \max_{a \in \text{dom}(A)} \Big\{ \max_{c \in \text{dom}(C)} \{ \min\{ R(A = a, B = b), R(B = b, C = c), \\
& \qquad\qquad R(A = a \mid A = a_{\text{obs}}) \} \} \Big\} \\
= \; & \max_{a \in \text{dom}(A)} \Big\{ \min\{ R(A = a, B = b), R(A = a \mid A = a_{\text{obs}}), \\
& \qquad \underbrace{\max_{c \in \text{dom}(C)} \{ R(B = b, C = c) \}}_{= R(B=b) \geq R(A=a, B=b)} \} \Big\} \\
= \; & \max_{a \in \text{dom}(A)} \{ \min\{ R(A = a, B = b), R(A = a \mid A = a_{\text{obs}}) \} \}.
\end{aligned}
$$

A:	color
B:	shape
C:	size

(1) holds because of the second axiom a discrete possibility measure has to satisfy.

(3) holds because of the fact that the relation R_{ABC} can be decomposed w.r.t. the set $\mathcal{M} = \{\{A, B\}, \{B, C\}\}$. ($A$: color, B: shape, C: size)

(2) holds, since in the first place

$$\begin{aligned} R(A = a, B = b, C = c | A = a_{obs}) &= R(A = a, B = b, C = c, A = a_{obs}) \\ &= \begin{cases} R(A = a, B = b, C = c), & \text{if } a = a_{\text{obs}}, \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

and secondly

$$\begin{aligned} R(A = a \mid A = a_{\text{obs}}) &= R(A = a, A = a_{\text{obs}}) \\ &= \begin{cases} R(A = a), & \text{if } a = a_{\text{obs}}, \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

and therefore, since trivially $R(A = a) \geq R(A = a, B = b, C = c)$,

$$\begin{aligned} & R(A = a, B = b, C = c \mid A = a_{obs}) \\ &= \min\{R(A = a, B = b, C = c), R(A = a \mid A = a_{\text{obs}})\}. \end{aligned}$$

Example 4:Relational Evidence Propagation, Step 2

A:	color
B:	shape
C:	size

$$
\begin{aligned}
&R(C = c \mid A = a_{\text{obs}})\\
&= R\Big(\bigvee_{a\in\text{dom}(A)} A = a, \bigvee_{b\in\text{dom}(B)} B = b, C = c \,\Big|\, A = a_{\text{obs}}\Big)\\
&\overset{(1)}{=} \max_{a\in\text{dom}(A)}\{\max_{b\in\text{dom}(B)}\{R(A = a, B = b, C = c \mid A = a_{\text{obs}})\}\}\\
&\overset{(2)}{=} \max_{a\in\text{dom}(A)}\{\max_{b\in\text{dom}(B)}\{\min\{R(A = a, B = b, C = c), R(A = a \mid A = a_{\text{obs}})\}\}\}\\
&\overset{(3)}{=} \max_{a\in\text{dom}(A)}\{\max_{b\in\text{dom}(B)}\{\min\{R(A = a, B = b), R(B = b, C = c),\\
&\qquad\qquad R(A = a \mid A = a_{\text{obs}})\}\}\}\\
&= \max_{b\in\text{dom}(B)}\{\min\{R(B = b, C = c),\\
&\qquad \underbrace{\max_{a\in\text{dom}(A)}\{\min\{R(A = a, B = b), R(A = a \mid A = a_{\text{obs}})\}\}\}}_{=R(B=b|A=a_{\text{obs}})}\\
&= \max_{b\in\text{dom}(B)}\{\min\{R(B = b, C = c), R(B = b \mid A = a_{\text{obs}})\}\}.
\end{aligned}
$$

Real World Example (continued)

Property family	Car body	Motor	Radio	Doors	Seat cover	Makeup mirrow	…
Property	Hatch-back	2.8 L 150 kW Otto	Type alpha	4	Leather, Type L3	yes	…

About 200 variables

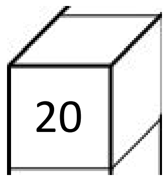
Typically 4 to 8, but up to 150 instances per variable

More than 2^{200} possible combinations available, for each combination an **installation rate** is needed.

The installation rate can be interpreted as a (subjective) **probability**.

Prior Probability

Cubes have a „probability“, e.g. 20/1000



all numbers in parts per 1000

Color marginals:

blue	red	yellow	green
220	330	170	280

large (300):

	blue	red	yellow	green	
△	20	90	10	80	400
□	2	1	20	17	240
○	28	24	5	3	360

medium (460):

	blue	red	yellow	green
△	18	81	9	72
□	8	4	80	68
○	56	48	10	6

small (240):

	blue	red	yellow	green
△	2	9	1	8
□	2	1	20	17
○	84	72	15	9

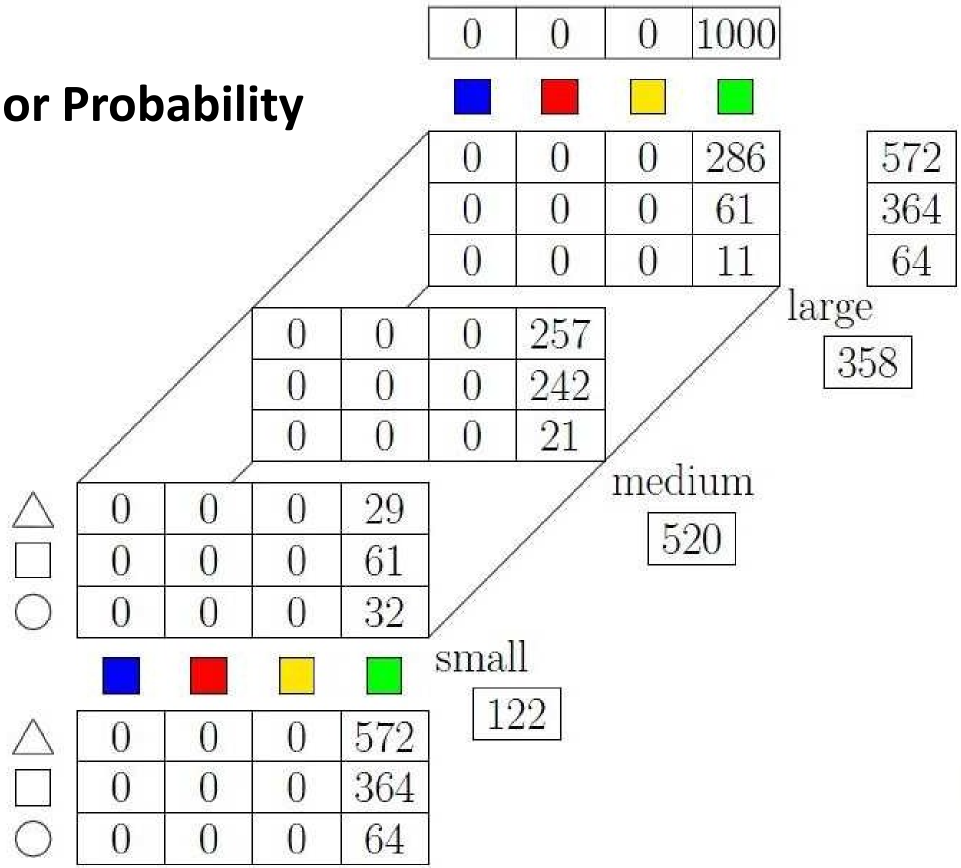
	blue	red	yellow	green
△	40	180	20	160
□	12	6	120	102
○	168	144	30	18

	s	m	l
△	20	180	200
□	40	160	40
○	180	120	60

	blue	red	yellow	green
large	50	115	35	100
medium	82	133	99	146
small	88	82	36	34

The numbers state the probability of the corresponding value combination. Compared to the example relation, the possible combinations are now frequent.

Posterior Probability



all numbers in parts per 1000

	s	m	l
△	29	257	286
□	61	242	61
○	32	21	11

large	0	0	0	358
medium	0	0	0	531
small	0	0	0	111

The concept is extremly simple: We have the evidence, that the given object is green. We calculate the conditional probability. Due to a normalization Color = Green has the „posterior" probability of 1.

For real applications the calculations are very complex. Decomposition helps to store and to update the probabilities in real applications.

Example 5: Probabilistic Decomposition

- As for relational networks, the three-dimensional probability distribution can be decomposed into projections to subspaces, namely the marginal distribution on the subspace formed by color and shape and the marginal distribution on the subspace formed by shape and size.
- The original probability distribution can be reconstructed from the marginal distributions using the following formulae $\forall i, j, k$:

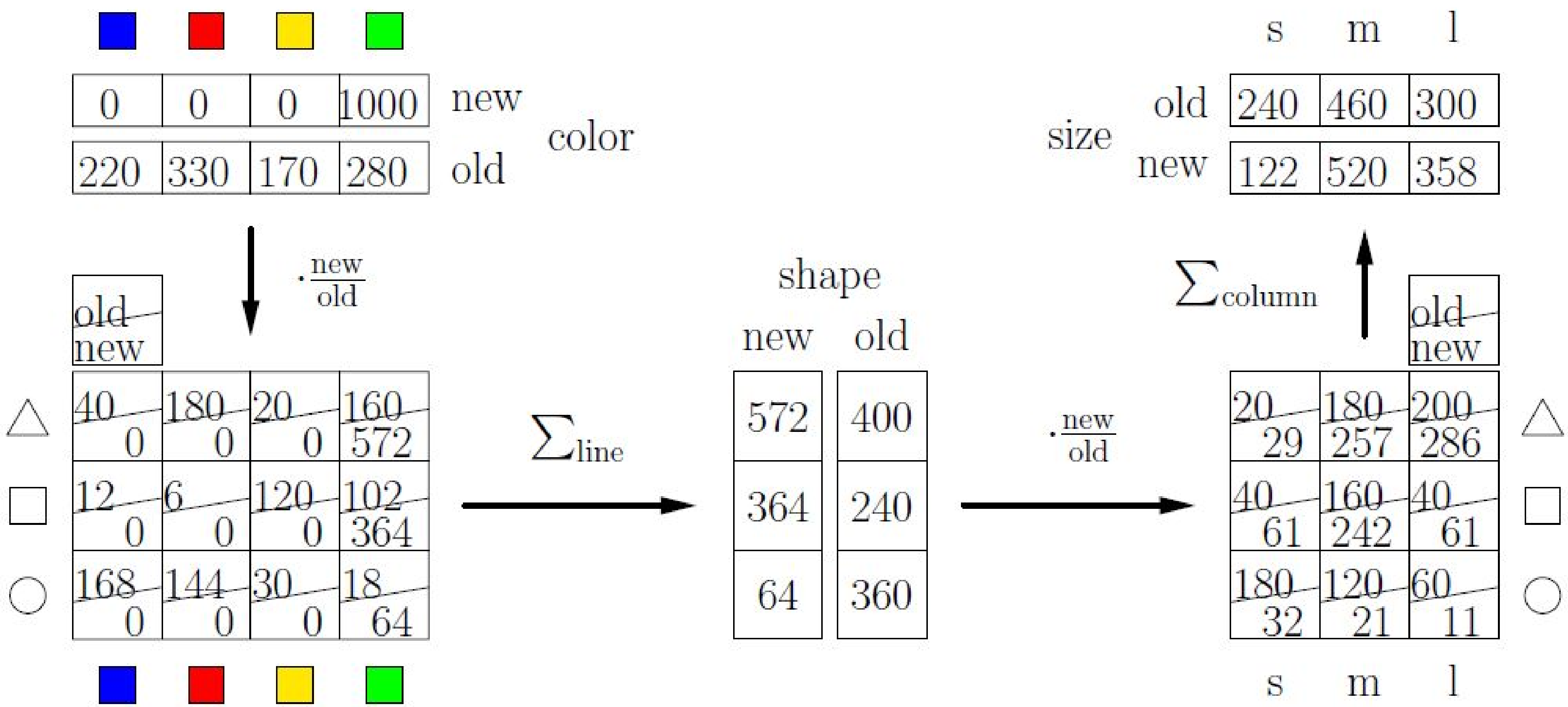
$$\begin{aligned} P(\omega_i^{(\text{color})}, \omega_j^{(\text{shape})}, \omega_k^{(\text{size})}) &= P(\omega_i^{(\text{color})}, \omega_j^{(\text{shape})}) \cdot P(\omega_k^{(\text{size})} \mid \omega_j^{(\text{shape})}) \\ &= P(\omega_i^{(\text{color})}, \omega_j^{(\text{shape})}) \cdot \frac{P(\omega_j^{(\text{shape})}, \omega_k^{(\text{size})})}{P(\omega_j^{(\text{shape})})} \end{aligned}$$

- These equations express the *conditional independence* of attributes *color* and *size* given the attribute *shape*, since they only hold if $\forall i, j, k$:

$$P(\omega_k^{(\text{size})} \mid \omega_j^{(\text{shape})}) = P(\omega_k^{(\text{size})} \mid \omega_i^{(\text{color})}, \omega_j^{(\text{shape})})$$

Example 5:Reasoning with Projections

Due to the fact that color and size are conditionally independent given the shape, the reasoning result can be obtained using only the projections to the subspaces:



This reasoning scheme can be formally justified with probability measures.

Example 5:Probabilistic Evidence Propagation, Step 1

A:	color
B:	shape
C:	size

$$
\begin{aligned}
&P(B = b \mid A = a_{\text{obs}}) \\
&= P\left(\bigvee_{a \in \text{dom}(A)} A = a, B = b, \bigvee_{c \in \text{dom}(C)} C = c \,\middle|\, A = a_{\text{obs}} \right) \\
&\overset{(1)}{=} \sum_{a \in \text{dom}(A)} \sum_{c \in \text{dom}(C)} P(A = a, B = b, C = c \mid A = a_{\text{obs}}) \\
&\overset{(2)}{=} \sum_{a \in \text{dom}(A)} \sum_{c \in \text{dom}(C)} P(A = a, B = b, C = c) \cdot \frac{P(A = a \mid A = a_{\text{obs}})}{P(A = a)} \\
&\overset{(3)}{=} \sum_{a \in \text{dom}(A)} \sum_{c \in \text{dom}(C)} \frac{P(A = a, B = b) P(B = b, C = c)}{P(B = b)} \cdot \frac{P(A = a \mid A = a_{\text{obs}})}{P(A = a)} \\
&= \sum_{a \in \text{dom}(A)} P(A = a, B = b) \cdot \frac{P(A = a \mid A = a_{\text{obs}})}{P(A = a)} \underbrace{\sum_{c \in \text{dom}(C)} P(C = c \mid B = b)}_{=1} \\
&= \sum_{a \in \text{dom}(A)} P(A = a, B = b) \cdot \frac{P(A = a \mid A = a_{\text{obs}})}{P(A = a)}.
\end{aligned}
$$

Example 5: Probabilistic Evidence Propagation, Step 1 (continued)

(1) holds because of Kolmogorov's axioms.

(3) holds because of the fact that the distribution p_{ABC} can be decomposed w.r.t. the set $\mathcal{M} = \{\{A, B\}, \{B, C\}\}$. ($A$: color, B: shape, C: size)

(2) holds, since in the first place

$$P(A = a, B = b, C = c \,|\, A = a_{obs}) = \frac{P(A = a, B = b, C = c, A = a_{obs})}{P(A = a_{\text{obs}})}$$

$$= \begin{cases} \dfrac{P(A = a, B = b, C = c)}{P(A = a_{\text{obs}})}, & \text{if } a = a_{\text{obs}}, \\ 0, & \text{otherwise,} \end{cases}$$

and secondly

$$P(A = a, A = a_{\text{obs}}) = \begin{cases} P(A = a), & \text{if } a = a_{\text{obs}}, \\ 0, & \text{otherwise,} \end{cases}$$

and therefore

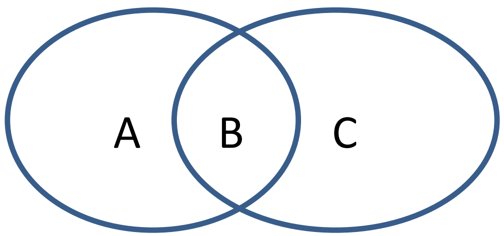
$$P(A = a, B = b, C = c \mid A = a_{obs})$$
$$= P(A = a, B = b, C = c) \cdot \frac{P(A = a \mid A = a_{\text{obs}})}{P(A = a)}.$$

Example 5:Probabilistic Evidence Propagation, Step 2

A:	color
B:	shape
C:	size

$$
\begin{aligned}
&P(C = c \mid A = a_{\text{obs}}) \\
&= P\left(\bigvee_{a \in \text{dom}(A)} A = a, \bigvee_{b \in \text{dom}(B)} B = b, C = c \,\middle|\, A = a_{\text{obs}} \right) \\
&\overset{(1)}{=} \sum_{a \in \text{dom}(A)} \sum_{b \in \text{dom}(B)} P(A = a, B = b, C = c \mid A = a_{\text{obs}}) \\
&\overset{(2)}{=} \sum_{a \in \text{dom}(A)} \sum_{b \in \text{dom}(B)} P(A = a, B = b, C = c) \cdot \frac{P(A = a \mid A = a_{\text{obs}})}{P(A = a)} \\
&\overset{(3)}{=} \sum_{a \in \text{dom}(A)} \sum_{b \in \text{dom}(B)} \frac{P(A = a, B = b) P(B = b, C = c)}{P(B = b)} \cdot \frac{P(A = a \mid A = a_{\text{obs}})}{P(A = a)} \\
&= \sum_{b \in \text{dom}(B)} \frac{P(B = b, C = c)}{P(B = b)} \underbrace{\sum_{a \in \text{dom}(A)} P(A = a, B = b) \cdot \frac{R(A = a \mid A = a_{\text{obs}})}{P(A = a)}}_{= P(B = b \mid A = a_{\text{obs}})} \\
&= \sum_{b \in \text{dom}(B)} P(B = b, C = c) \cdot \frac{P(B = b \mid A = a_{\text{obs}})}{P(B = b)}.
\end{aligned}
$$

Decomposition in Subspaces



Subspace (A,B) Subspace(B,C)

$$P(A,B,C) = P(A,B)P(B,C)/P(B)$$

Markov Network

Decomposition using Dependencies

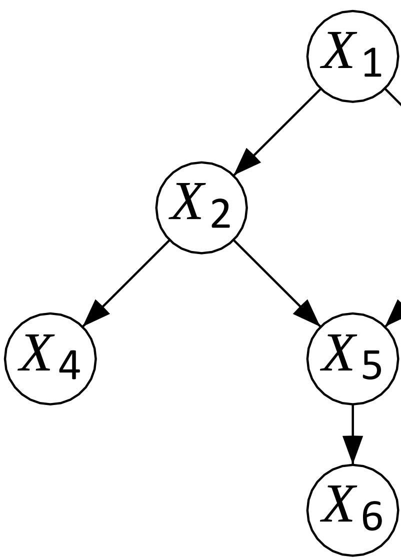
$$A \rightarrow B \rightarrow C$$

$$P(A,B,C) = P(A)P(B|A)P(C|B)$$

Bayes Network

Ecample 6: Bayesian Network

Bayes Networks are directed acyclic graphs (DAGs) where the nodes represent random variables. For each node X, the conditional probability of X with respect to its direct predecessors (the „father“ nodes) is calculated. The common probability of all nodes is defined as the product of the conditional probabilites.

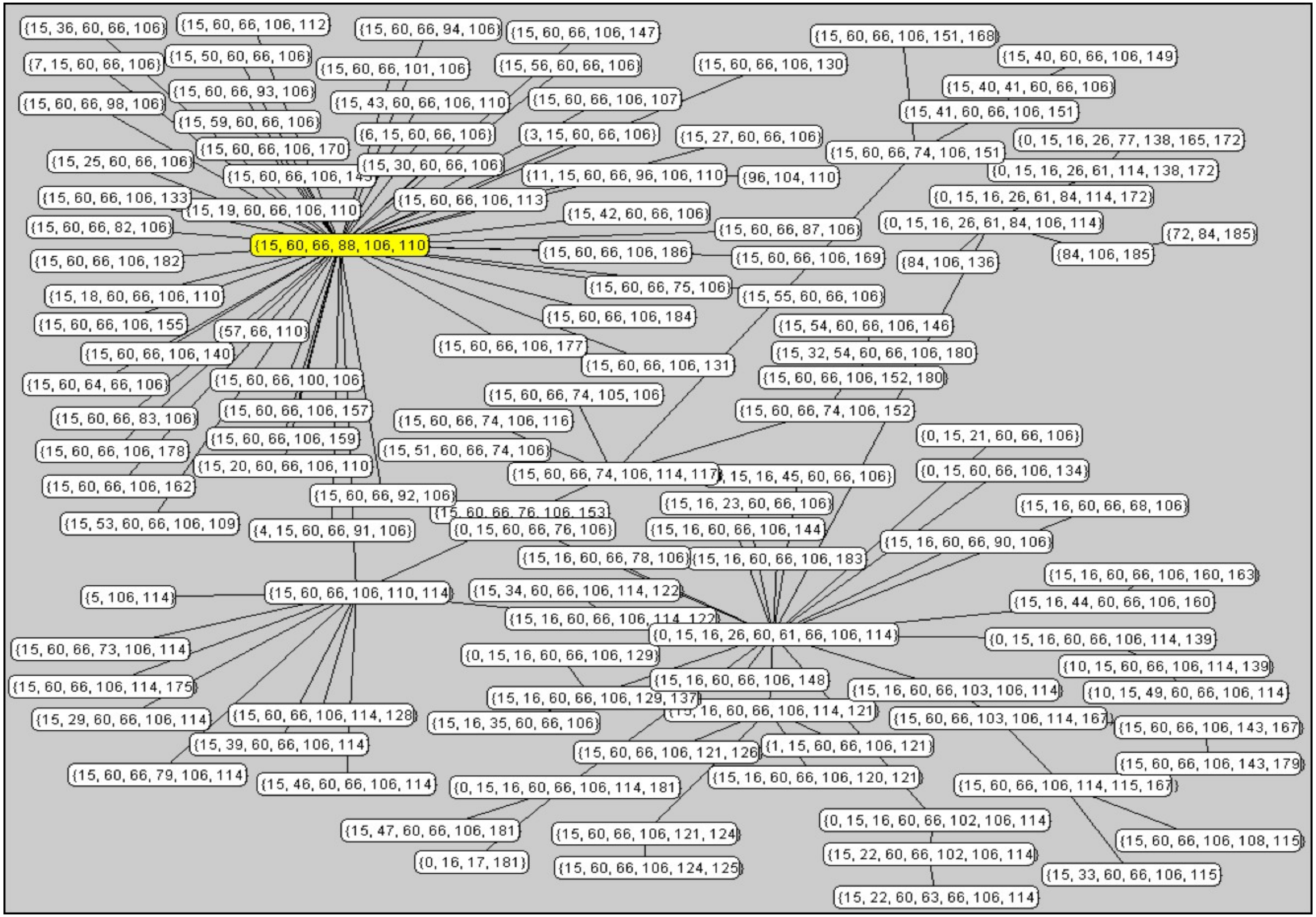


Given a DAG, we define the probability according to the (in)dependency structure:

$$
\begin{aligned}
P(X_1, \ldots, X_6) = \; & P(X_6 \mid X_5) \cdot \\
& P(X_5 \mid X_2, X_3) \cdot \\
& P(X_4 \mid X_2) \cdot \\
& P(X_3 \mid X_1) \cdot \\
& P(X_2 \mid X_1) \cdot \\
& P(X_1)
\end{aligned}
$$

Real World Example (continues) Markov Net

Property Families for VW Bora



Each number corresponds to an attribute.

The 186 attributes have 2 to 20 different values.

Using the installation rates we obtain a 186 dimensional probability space.

This high dimensional probability space stored by decomposing it by using 174 low dimensional marginal probability spaces.

How to calculate conditional probabilities?

Probabilistic Decomposition

Definition: Let $U = \{A_1, \ldots, A_n\}$ be a set of attributes and p_U a probability distribution over U. Furthermore, let $\mathcal{M} = \{M_1, \ldots, M_m\} \subseteq 2^U$ be a set of nonempty (but not necessarily disjoint) subsets of U satisfying

$$\bigcup_{M \in \mathcal{M}} M = U.$$

p_U is called **decomposable** or **factorizable** w.r.t. $\mathcal{M}$ iff it can be written as a product of m nonnegative functions $\phi_M : \mathcal{E}_M \to \mathbb{R}_0^+$, $M \in \mathcal{M}$, i.e., iff

$$\forall a_1 \in \mathrm{dom}(A_1) : \ldots \forall a_n \in \mathrm{dom}(A_n) :$$

$$p_U\left(\bigwedge_{A_i \in U} A_i = a_i\right) = \prod_{M \in \mathcal{M}} \phi_M\left(\bigwedge_{A_i \in M} A_i = a_i\right).$$

If p_U is decomposable w.r.t. $\mathcal{M}$ the set of functions

$$\Phi_{\mathcal{M}} = \{\phi_{M_1}, \ldots, \phi_{M_m}\} = \{\phi_M \mid M \in \mathcal{M}\}$$

is called the **decomposition** or the **factorization** of p_U.
The functions in $\Phi_{\mathcal{M}}$ are called the **factor potentials** of p_U.

Summary

It is often possible to exploit local constraints (wherever they may come from — both structural and expert knowledge-based) in a way that allows for a decomposition of the large (intractable) distribution $P(X_1, ..., X_n)$ into several sub-structures $\{C_1, ..., C_m\}$ such that:

- The collective size of those sub-structures is much smaller than that of the original distribution P.
- The original distribution P is decomposable (with no or at least as few as possible errors) from these sub-structures.
- This decomposition allows for efficient propagation algorithms for integration of new evidence.