Fuzzy Data Analysis

Possibilistic Reasoning

Prof. Dr. Rudolf Kruse

A Simple Example

Oil contamination of water by trading vessels

Typical formulation:

"The accident occurred 10 miles away from the coast."

Locations of interest: *open sea* (z3), *12-mile zone* (z2), *3-mile zone* (z1), *canal* (ca), *refueling dock* (rd), *loading dock* (ld)

These 6 locations Ω are disjoint and exhaustive

$$\Omega = \{z3, z2, z1, ca, rd, ld\}$$

Modeling Degrees of Belief

Statements are often not simply true or false. Decision maker should be able to quantify their „degree of belief". This can be an objective measurement or subjective valuation. The standard way to model such situations with uncertainty is to use probability theory:

Sample space Θ (finite set of distinct possible outcomes of some random experiment), Events of interest are subsets $A \subseteq \Theta$. The probabilities are subjectively interpreted as **degrees of belief.**

$P : 2^{\Theta} \rightarrow [0, 1]$ are then required to satisfy the Kolmogorov axioms. There are good arguments for using probabilities for, modelling beliefs, e.g. the so called „Dutch Book argument“.

Kolmogorov Axioms

For **finite** Θ, probability function $P : 2^{\Theta} \rightarrow [0, 1]$ must satisfy

i) $0 \leq P(A) \leq 1$ for all events $A \subseteq \Theta$,

ii) $P(\Theta) = 1$,

iii) if $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$ for all A,B

Simple Example

Consider the subjective statement: „The ship is in ca or rd or ld with degree of certainty 0.6, that‘s all I know.“

A modelling with probility theory forces the user to specify the degrees of belief for all elementary events. In the subjective statement above the expert did not want to that.

An option could be to assign the probabilities
P(ca)=P(rd)=P(ca) =0.2, P(z1)=P(z2)=P(z3)= 0.4/3
This is a very precise (too precise) information that doesn‘t reflect the state of the knowledge - namely the partial ignorance of the expert.

An alternative solution is to assign beliefs directly to subsets and not to elements, so called mass distributions.

Mass Distribution

Recall example with $\Omega = \{z3, z2, z1, ca, rd, ld\}$

Propositional statement *in port* equals event $\{ca, rd, ld\}$

Event may represent maximum level of differentiation for expert

Expert specifies **mass distribution** $m : 2^{\Omega} \to [0, 1]$

Here, Ω is called **frame of discernment**

$m : 2^{\Omega} \to [0, 1]$ must satisfy

(i) $m(\emptyset) = 0$,

(ii) $\sum_{A:A \subseteq \Omega}(A) = 1$

Subsets $A \subseteq \Omega$ with $m(A) > 0$ are called **focal elements** of m

Belief and Plausibility

$m(A)$ measures belief committed *exactly* to A

For *total* amount of belief (**credibility**) of A, sum up $m(B)$ whereas $B \subseteq A$

For *maximum* amount of belief movable to A, sum up $m(B)$ with $B \cap A \neq \emptyset$ **(plausibility)**

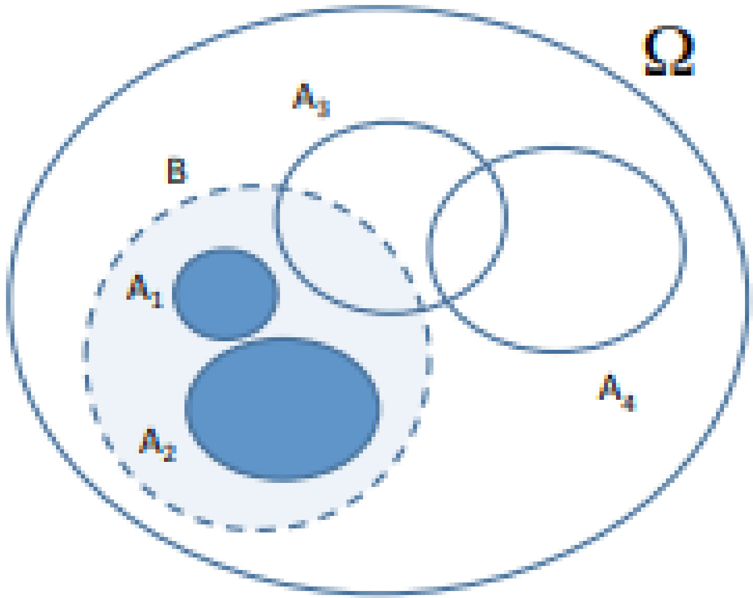
This leads to **belief function** and **plausibility function**

$$\mathrm{Bel}_m : \ 2^{\Omega} \to [0,1], \quad \mathrm{Bel}_m(A) = \sum_{B : B \subseteq A} m(B)$$

$$\mathrm{Pl}_m : \ 2^{\Omega} \to [0,1], \quad \mathrm{Pl}_m(A) = \sum_{B : B \cap A \neq \emptyset} m(B)$$

Belief and Plausibility

If the evidence tells us that the truth is in A, and $A \subseteq B$, we say that the evidence supports B.



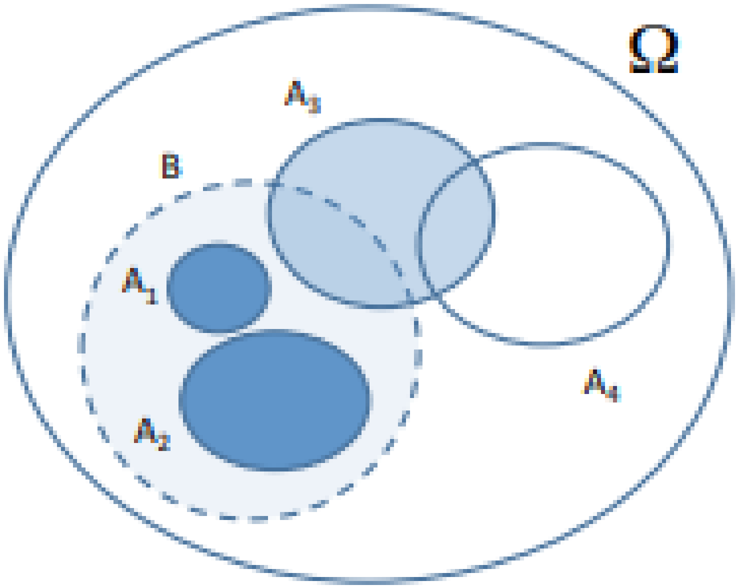
- Given a normalized mass function m, the probability that the evidence supports B is thus

$$Bel(B) = \sum_{A \subseteq B} m(A)$$

- The number $Bel(B)$ is called the degree of belief in B, and the function $B \to Bel(B)$ is called a belief function.

Belief and Plausibility

If the evidence does not support $\overline{B}$, it is consistent with B.



- The probability that the evidence is consistent with B is thus

$$Pl(B) = \sum_{A \cap B \neq \emptyset} m(A)$$
$$= 1 - Bel(\overline{B}).$$

- The number $Pl(B)$ is called the plausibility of B, and the function $B \to Pl(B)$ is called a plausibility function.

Example

Consider statement: “ship is *in port* with degree of certainty of 0.6, further evidence is not available”

Mass distribution
$m : 2^{\Omega} \to [0,1], m(\{\text{in port}\}) = 0.6, m(\Omega) = 0.4, m(A) = 0$ otherwise

$m(\Omega) = 0.4$ represents inability to attach that amount of mass to any A, which is different from Ω

e.g. $m(\{\overline{\text{in port}}\}) = 0.4$ would exceed expert's statement

Properties of Belief Functions

Function $Bel : 2^\Omega \to [0,1]$ is a completely monotone capacity: it verifies $Bel(\emptyset) = 0$, $Bel(\Omega) = 1$ and

$$Bel\left(\bigcup_{i=1}^{k} A_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1} Bel\left(\bigcap_{i \in I} A_i\right).$$

for any $k \geq 2$ and for any family $A_1, \ldots, A_k$ in 2^Ω.

Conversely, to any completely monotone capacity Bel corresponds a unique mass function m such that:

$$m(A) = \sum_{\emptyset \neq B \subseteq A} (-1)^{|A|-|B|} Bel(B), \quad \forall A \subseteq \Omega.$$

Relations between m, Bel, and PL

Let m be a mass function, Bel and Pl the corresponding belief and plausibility functions

For all $A \subseteq \Omega$,

$$Bel(A) = 1 - Pl(\overline{A})$$

$$m(A) = \sum_{\emptyset \neq B \subseteq A} (-1)^{|A|-|B|} Bel(B)$$

$$m(A) = \sum_{B \subseteq A} (-1)^{|A|-|B|+1} Pl(\overline{B})$$

m, Bel and Pl are thus three equivalent representations of

- a piece of evidence or, equivalently
- a state of belief induced by this evidence

Belief and Plausibility

In any case $\text{Bel}(\Omega) = 1$ ("closed world" assumption)

Total ignorance modeled by $m_0 : 2^\Omega \to [0,1]$ with $m_0(\Omega) = 1$, $m_0(A) = 0$ for all $A \neq \Omega$

m_0 leads to $\text{Bel}(\Omega) = \text{Pl}(\Omega) = 1$ and $\text{Bel}(A) = 0, \text{Pl}(A) = 1$ for all $A \neq \Omega$

For ordinary probability, use $m_1 : 2^\Omega \to [0,1]$ with $m_1(\{\omega\}) = p_\omega$ and $m_1(A) = 0$ for all sets A with $|A| > 1$

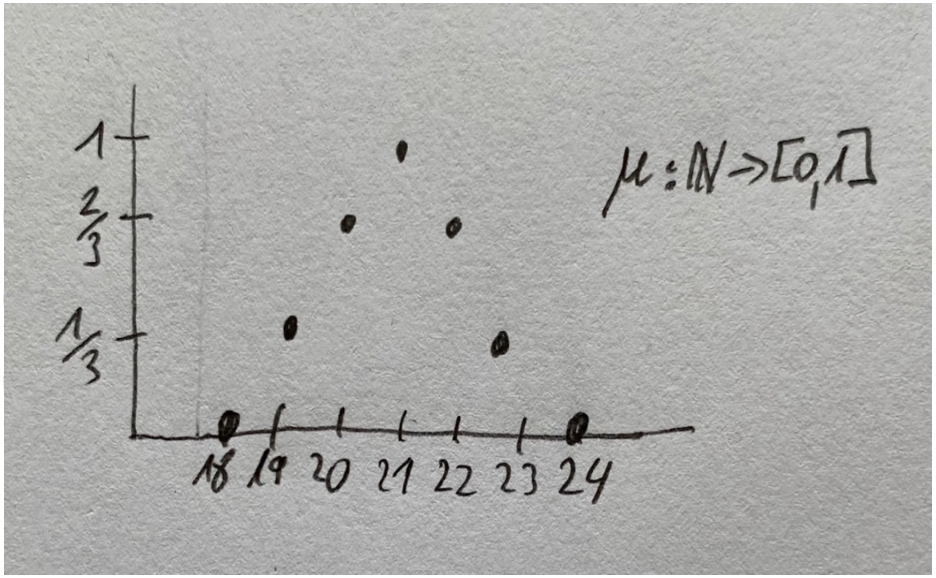
m_1 is called Bayesian belief function

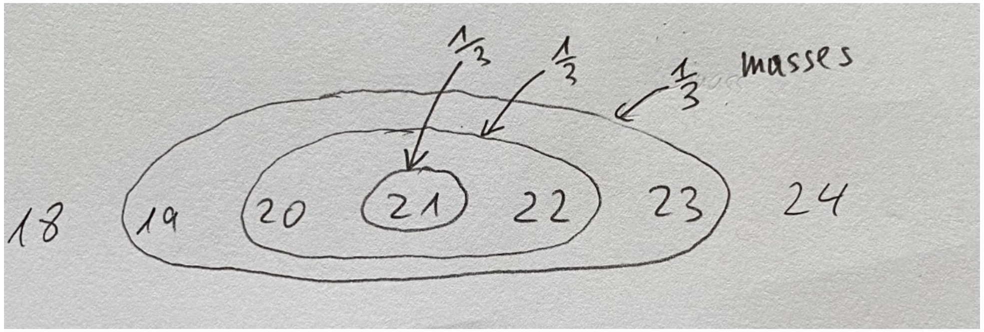
Exact knowledge modeled by $m_2 : 2^\Omega \to [0,1]$, $m_2(\{\omega_0\}) = 1$ and $m_2(A) = 0$ for all $A \neq \{\omega_0\}$

Fuzzy Sets induce Plausibility Measures

Let variable T be temperature in $^\circ C$ (only integers)

Current but unknown value T_0 is given by "T is around $21^\circ C$"





Suppose a fuzzy set μ is normal and has a finite number of different membership degrees.
Then μ induces a plausibility measure Pl by
Pl({x})= μ(x) , for all x , and Pl(B) = max{ Pl(x) : x ∈ B } for all B.
Let m be the corresponding mass assignment m. Its focal element (i.e. the subsets with positive mass) are nested : A1 subset of A2, A2 subset of A3, etc). The focal elements are the alpha cuts of μ.

Possibility and Necessity Measures

These ideas can be expressed in a simpler way by using possibility measures:
We describe an imprecise value by giving possibility degrees to all values. It would be strange, if we consider **no** value as possible, so we say that a possibility distribution is a function $\pi : X \to [0,1]$ if there is at least one x with $\pi(x) = 1$

The corresponding **possibility measure is defined by**

$$\Pi : 2^{\Omega} \to [0,1], \qquad \Pi(B) = \max\{\pi(\omega) : \omega \in B\}$$

Properties of Possibility Measures

i) $\Pi(\emptyset) = 0$

ii) $\Pi(\Omega) = 1$

iii) $\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$ for all $A, B \subseteq \Omega$

Possibility of some set is determined by its "most possible" element

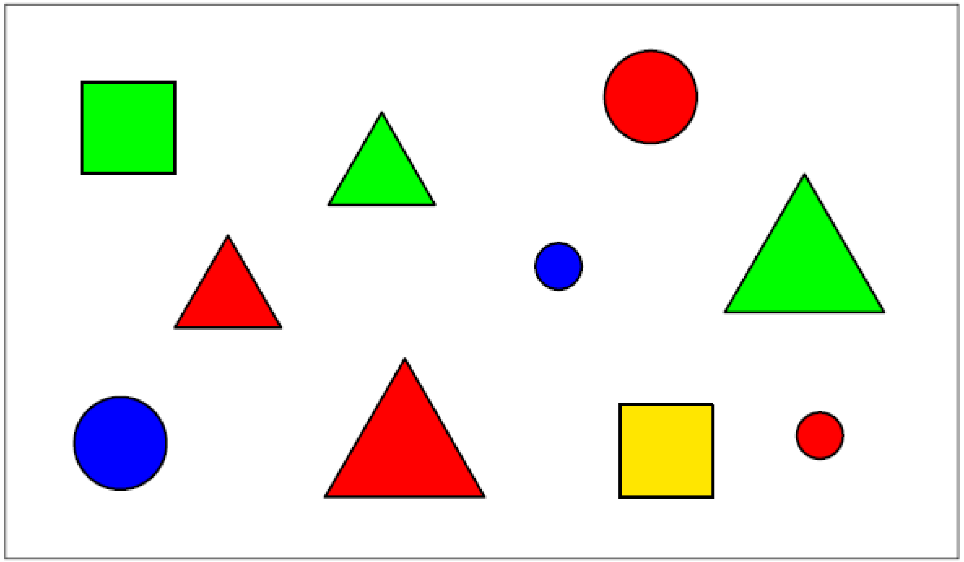
$\text{nec}(\Omega) = 1 - \Pi(\emptyset) = 1$ means closed world assumption:
"necessarily $\omega_0 \in \Omega$" must be true

Total ignorance: $\Pi(B) = 1, \text{nec}(B) = 0$ for all $B \neq \emptyset, B \neq \Omega$

Perfect knowledge: $\Pi(\{\omega\}) = \text{nec}(\{\omega\}) = 0$ for all $\omega \neq \omega_0$ and
$\Pi(\{\omega_0\}) = \text{nec}(\{\omega_0\}) = 1$

Example

Example Domain

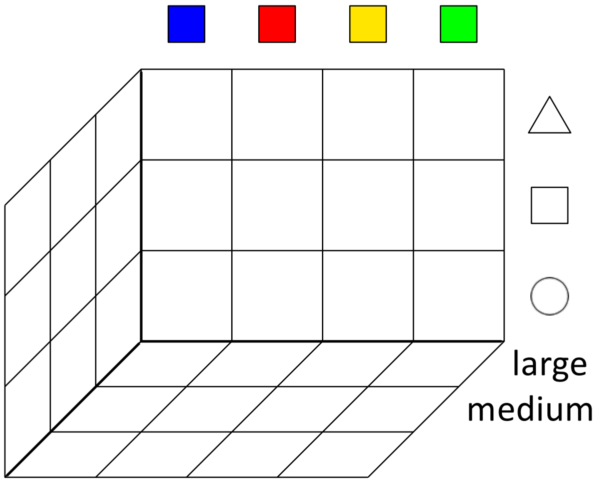


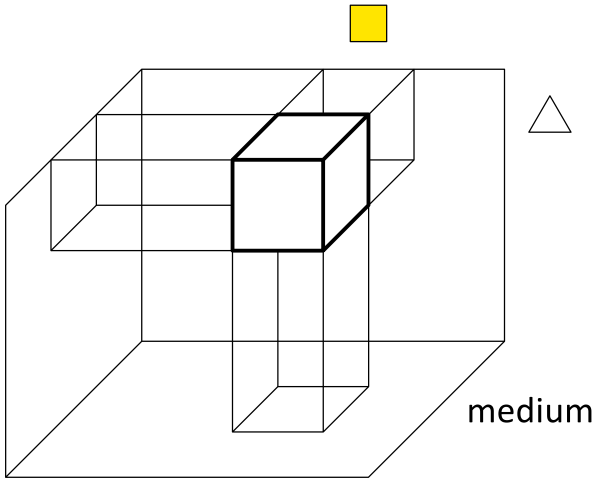
Relation

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

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

Example: Representation as a Relation





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.

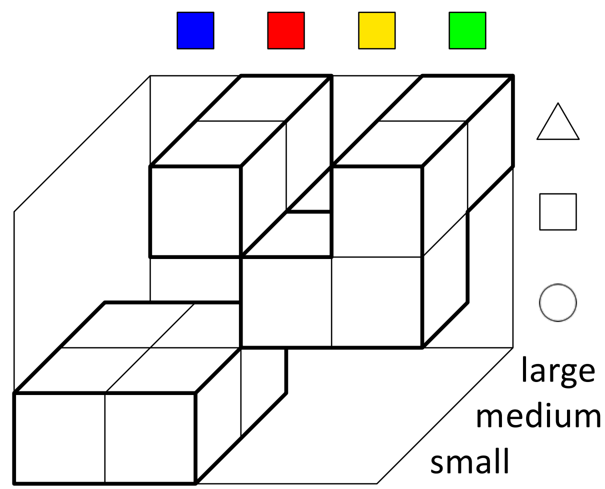
The events in a domain are mutually exclusive and exhaustive.

Example: Relation in a many-dimensional space

Relation

color	shape	size
blue ■	○	small
blue ■	○	medium
red ■	○	small
red ■	○	medium
red ■	△	medium
red ■	△	large
yellow ■	□	medium
green ■	□	medium
green ■	△	medium
green ■	△	large

Visual Description



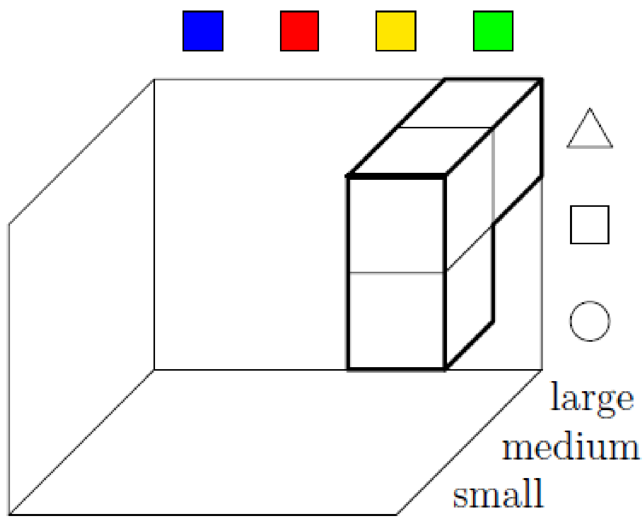
Each cube represents one tuple.

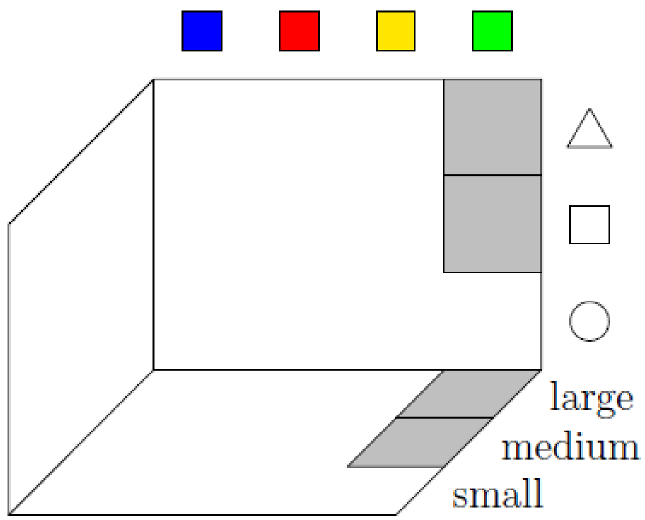
The spatial representation helps to understand the decomposition mechanism.

Example: Reasoning

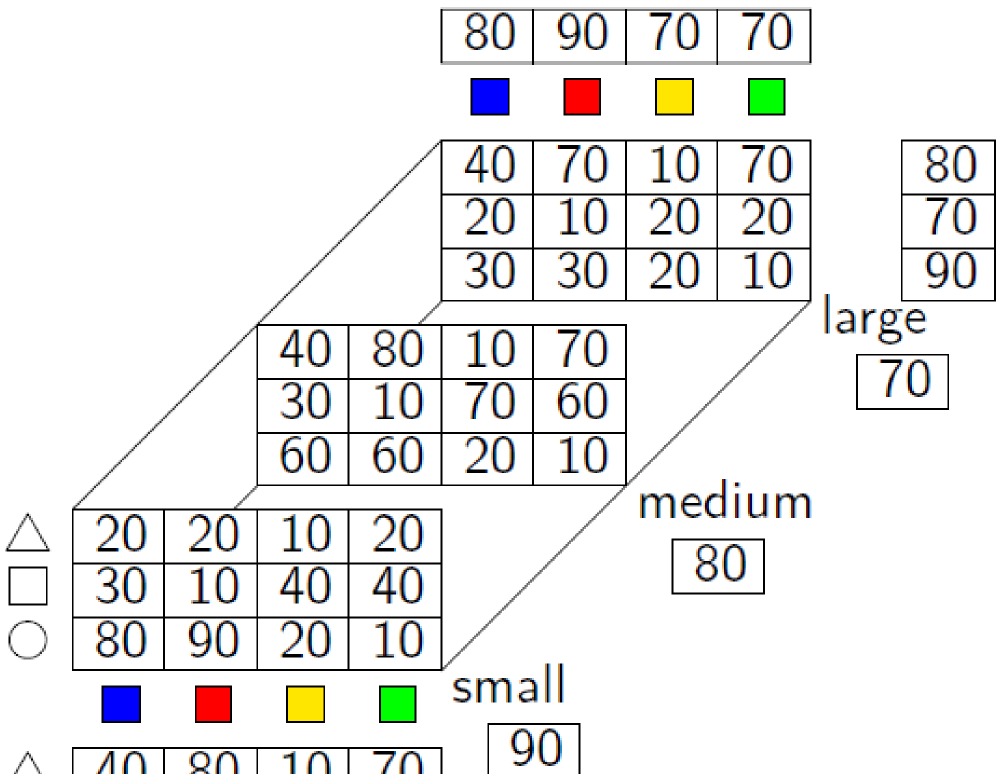
- Let it be known (e.g. from an observation) that the given object is green.
  This information considerably reduces the space of possible value combinations.

- From the prior knowledge it follows that the given object must be
  - either a triangle or a square and
  - either medium or large.





Example: Extension to possibility distributions



△	40	80	10	70
□	30	10	70	60
○	80	90	20	10

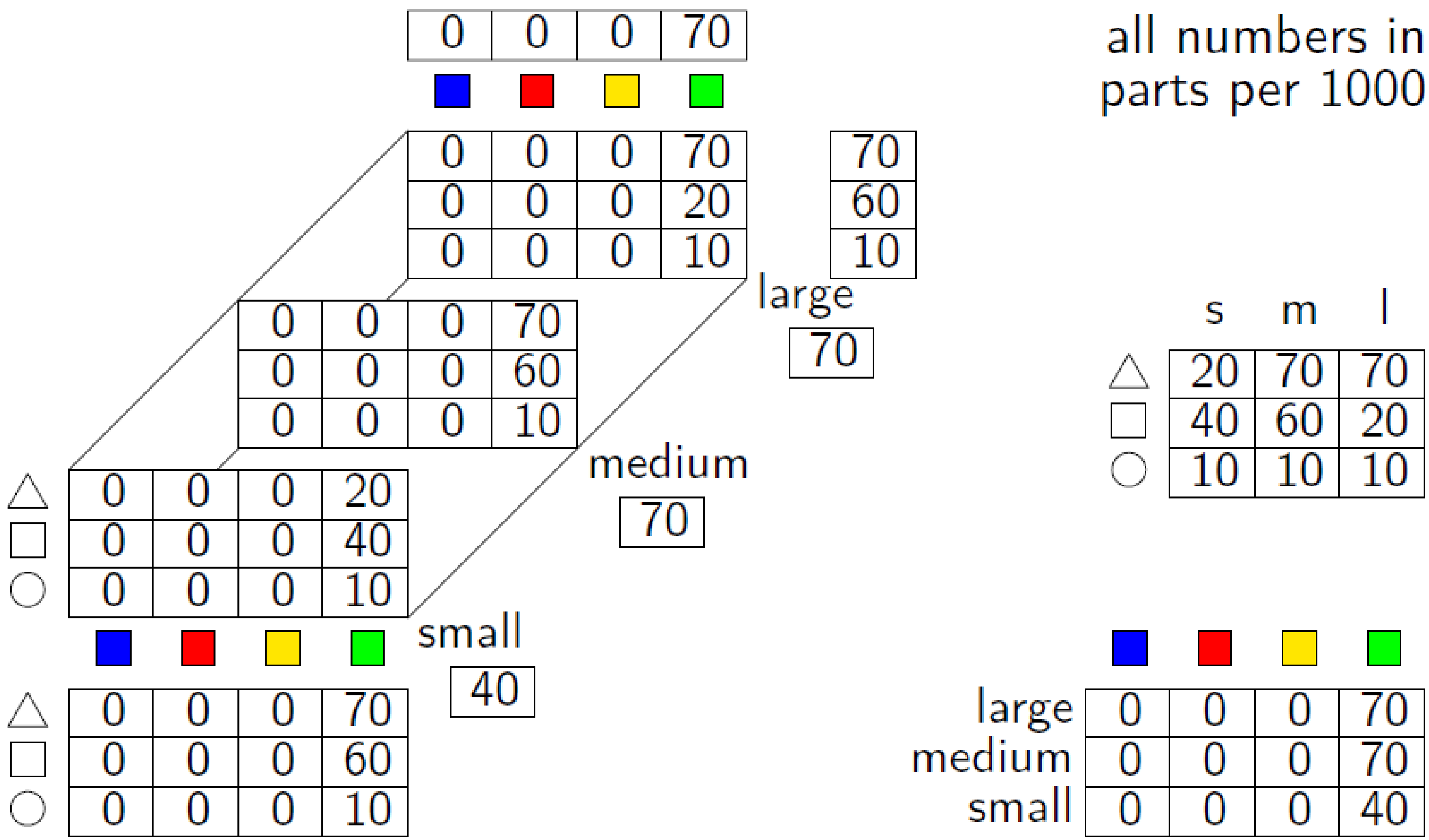
all numbers in parts per 1000

	s	m	l
△	20	80	70
□	40	70	20
○	90	60	30

large	40	70	20	70
medium	60	80	70	70
small	80	90	40	40

Numbers state degrees of possibility of corresponding value combination

Example: Reasoning



From the information, that the object is green, we can derive information obout the possibilities of shape and size. For high dimensional possilities the complexity can be handled by using information about (conditional) independences

Possibilistic Networks

Example: Decomposition of a 21-dim possibility distribution by using independences between lower dimensional possibility distributions .

The (hyper-) graph visualized the independence structure by separation properties in the graph, and this representation allows efficient reasoning and learning methods in high dimensional problems.