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 $\Omega$ are disjoint and exhaustive

$$\Omega = \{z_3, z_2, z_1, ca, rd, ld\}$$
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:

\[ \Omega \] (finite set of distinct possible outcomes of some random experiment), Events of interest are subsets \( A \subseteq \Omega \). The probabilities are subjectively interpreted as degrees of belief.

\[ P : 2^\Omega \rightarrow [0, 1] \] are then required to satisfy the Kolmogorov axioms. There are good arguments for using probabilities for modeling beliefs, e.g. the so called "Dutch Book argument".
Kolmogorov Axioms

For finite $\Theta$, 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 $A, B$
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 \( P(ca) = P(rd) = P(ld) = 0.2, P(z1) = P(z2) = P(z3) = 0.4/3 \). 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.
Recall example with \( \Omega = \{z_3, z_2, z_1, ca, rd, ld\} \)

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

Event may represent maximum level of differentiation for expert

Expert specifies \textbf{mass distribution} \( m : 2^\Omega \rightarrow [0, 1] \)

Here, \( \Omega \) is called \textbf{frame of discernment}

\[ m : 2^\Omega \rightarrow [0, 1] \text{ must satisfy} \]
\begin{enumerate}
  
  \item (i) \( m(\emptyset) = 0 \),
  
  \item (ii) \( \sum_{A : A \subseteq \Omega} (A) = 1 \)
\end{enumerate}

Subsets \( A \subseteq \Omega \) with \( m(A) > 0 \) are called \textbf{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

$$Bel_m : 2^\Omega \rightarrow [0, 1], \quad Bel_m(A) = \sum_{B: B \subseteq A} m(B)$$

$$Pl_m : 2^\Omega \rightarrow [0, 1], \quad Pl_m(A) = \sum_{B: B \cap A \neq \emptyset} m(B)$$
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 $B \rightarrow Bel(B)$ is called a belief function.
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 $B \rightarrow PL(B)$ is called a **plausibility function**.
Consider statement: “ship is in port with degree of certainty of 0.6, further evidence is not available”

Mass distribution

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

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

e.g. \( m(\{\text{in port}\}) = 0.4 \) would exceed expert’s statement
Function $Bel : 2^\Omega \rightarrow [0, 1]$ is a \textit{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 $2^\Omega$.

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

$$m(A) = \sum_{\emptyset \neq B \subseteq A} (-1)^{|A|-|B|} Bel(B), \quad \forall A \subseteq \Omega.$$
Let $m$ be a mass function, $Bel$ and $Pl$ the corresponding belief and plausibility functions.

For all $A \subseteq \Omega$,

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

$$m(A) = \sum_{0 \neq B \subseteq A} (-1)^{|A| - |B|} Bel(B)$$

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

$m$, $Bel$ 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 $m_0 : 2^\Omega \rightarrow [0, 1]$ 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 $m_1 : 2^\Omega \rightarrow [0, 1]$ with $m_1(\{\omega\}) = \rho_\omega$ and $m_1(A) = 0$ for all sets $A$ with $|A| > 1$

$m_1$ is called Bayesian belief function

Exact knowledge modeled by $m_2 : 2^\Omega \rightarrow [0, 1]$, $m_2(\{\omega_0\}) = 1$ and $m_2(A) = 0$ for all $A \neq \{\omega_0\}$
Fuzzy Sets induce Plausibility Measures

Let variable $T$ be temperature in °C (only integers)

Current but unknown value $T_0$ is given by “$T$ is around 21° C”

Suppose a fuzzy set $\mu$ is normal and has a finite number of different membership degrees. Then $\mu$ induces a plausibility measure $\text{Pl}$ by

$$\text{Pl}(\{x\}) = \mu(x), \text{ for all } x,$$

and

$$\text{Pl}(B) = \max\{\text{Pl}(x) : x \in B\} \text{ for all } B.$$

Let $m$ be the corresponding mass assignment $m$. Its focal element (i.e. the subsets with positive mass) are nested: $A_1$ subset of $A_2$, $A_2$ subset of $A_3$, etc.). The focal elements are the alpha cuts of $\mu$. 
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

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

Example

- 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

<table>
<thead>
<tr>
<th>color</th>
<th>shape</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>◌</td>
<td>small</td>
</tr>
<tr>
<td>□</td>
<td>◌</td>
<td>medium</td>
</tr>
<tr>
<td>□</td>
<td>◌</td>
<td>small</td>
</tr>
<tr>
<td>□</td>
<td>◌</td>
<td>medium</td>
</tr>
<tr>
<td>□</td>
<td>△</td>
<td>medium</td>
</tr>
<tr>
<td>□</td>
<td>△</td>
<td>large</td>
</tr>
<tr>
<td>□</td>
<td>□</td>
<td>medium</td>
</tr>
<tr>
<td>□</td>
<td>□</td>
<td>medium</td>
</tr>
<tr>
<td>□</td>
<td>△</td>
<td>medium</td>
</tr>
<tr>
<td>□</td>
<td>△</td>
<td>large</td>
</tr>
</tbody>
</table>

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 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

Numbers state degrees of possibility of corresponding value combination.
Example: Reasoning

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