Reasoning with uncertainty

Carles Noguera i Clofent

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

---

An acknowledgment is due to Lluís Godo for his help in preparing these slides.
Uncertainty arises from the lack of information.

Uncertainty and imprecision (vagueness) are different forms of imperfect information, although they can be found together.
Outline

1. Probabilistic logic
2. Possibilistic logic
Basic probability notions – 1

\( S = \{s_1, \ldots, s_n\} \): finite set of states
\( \mathcal{P}(S) \): set of events (propositions)

A probability measure is a function

\[ P : \mathcal{P}(S) \to [0, 1] \]

such that

- \( P(\emptyset) = 0 \)
- \( P(S) = 1 \)
- \( P(A \cup B) = P(A) + P(B) - P(A \cap B) \)
A probability measure is characterized by its corresponding probability distribution,

\[ p : S \rightarrow [0, 1] \]

\[ p(s_i) \mapsto P\{s_i\} \]

\[ \sum_{i=1}^{n} p(s_i) = 1 \]
Basic probability notions – 3

\[ P(A) = 0 \text{ iff } A \text{ is impossible} \]

\[ P(A) = 1 \text{ iff } A \text{ is necessary} \]

\[ P(\neg A) = 1 - P(A) \]
Where do probabilities come from?

Why $p(s_i) = p_i$?

- **Objective probabilities:**
  1. $p_i = \frac{\text{favorable cases for } s_i}{\text{total number of possible cases}}$ (assuming equipossibility)
  2. $p_i = \text{limit of the relative frequency of } s_i \text{ in a large number of trials}$

- **Subjective probabilities:** $p_i = \text{degree of belief that } s_i \text{ is the actual state}$
  De Finetti: $p_i = \text{price a gambler is willing to pay when betting on } s_i \text{ with a banker that can reverse rôles}$
Conditional probability – 1

\[ P(A \mid B) = \text{probability of } A \text{ given } B, \text{ i.e. probability of } A \text{ when } B \text{ is known to occur} \]

\[ P(A \mid B) = \frac{P(A \cap B)}{P(B)} \text{ (whenever } P(B) > 0) \]

\( A \) and \( B \) are independent events iff \( P(A \mid B) = P(A) \) iff

\[ P(A \cap B) = P(A) \cdot P(B) \]

\( P_B : \mathcal{P}(S) \rightarrow [0, 1], \text{ given by } P_B(A) = P(A \mid B) \text{ is the induced probability measure} \)
Conditional probability – 2

\[ P(A \mid B) \] is not the probability of \( A \) if \( B \) is true.

Example

\[ P: \text{uniform probability on } \{1, 2, 3, 4, 5\} \]

\[ P(\text{Even} \mid \{1, 2, 3\}) = P(\text{Even} \mid \{3, 4, 5\}) = 1/3 \]

\[ C_1 \text{ if result } \in \{1, 2, 3\}, \text{ then } P(\text{Even}) = 1/3 \]

\[ C_2 \text{ if result } \in \{3, 4, 5\}, \text{ then } P(\text{Even}) = 1/3 \]

In classical logic, from \( C_1 \) and \( C_2 \), one obtains that
\[ P(\text{Even}) = 1/3! \text{ (absurd because } P(\text{Even}) = 2/5) \]
Bayesian networks

- Represent conditional dependence relations in a **directly acyclic graph** (DAG)
- Use the chain rule:
  \[
P(A \cap B \cap C) = P(A \mid B \cap C) \cdot P(B \mid C) \cdot P(C)
\]
- Simplify the expression with possible conditional independence relations
- (more information about Bayesian networks is available in previous lectures of this master by Josep Puyol)
A variety of formalisms that try to combine probability theory with mathematical logic in order to handle reasoning under uncertainty (Nilsson, Halpern, Hájek, etc.)

Knowledge base: $KB = \{\langle \text{logical_expression}, \alpha \rangle \}$

**Example**

- $\langle p \land q, 0.6 \rangle \quad P(p \land q) = 0.6$
- $\langle \neg p \lor q, 0.75 \rangle \quad P(\neg p \lor q) = 0.75$
- $\langle p \rightarrow q, 0.4 \rangle \quad P(q \mid p) = 0.4$
Each element of $KB$ gives a constraint on an unknown probability measure $P : S \rightarrow [0, 1]$

Inference: solving an optimization problem, maximize or minimize an algebraic expression (quotient of summations) under a set of (possibly non-linear) constraints.
Probabilistic logics: an example – 1

\[ KB = \{ \langle A \rightarrow D, p_1 \rangle, \langle \neg A \rightarrow D, p_2 \rangle, \langle B \rightarrow D, p_3 \rangle, \langle D \rightarrow B, p_4 \rangle, \langle A \rightarrow B, p_5 \rangle, \langle A, p_6 \rangle \} \]

Case: Symptoms \( \neg A \) and \( B \)

Diagnosis: \( D \)

Question: \( P(D \mid \neg A \land B) \)?
Probabilistic logic
Possibilistic logic

Probabilistic logics: an example – 2

Probabilistic interpretation of the knowledge base:

- \( p_1 = P(D \mid A) = \frac{P(A \land D)}{P(A)} \)
- \( p_2 = P(D \mid \neg A) = \frac{P(\neg A \land D)}{P(\neg A)} \)
- \( p_3 = P(D \mid B) = \frac{P(B \land D)}{P(B)} \)
- \( p_4 = P(B \mid D) = \frac{P(B \land D)}{P(D)} \)
- \( p_5 = P(B \mid A) = \frac{P(A \land B)}{P(A)} \)
- \( p_6 = P(A) \)

unknown: \( P(D \mid \neg A \land B) = \frac{P(\neg A \land B \land D)}{P(\neg A \land B)} \) ?

Possible worlds: \( \omega_i = (\neg)A \land (\neg)B \land (\neg)D \) (\( 2^3 = 8 \) worlds)
Possible worlds (states):

\[ \omega_1 = A \land B \land D \quad \omega_2 = A \land B \land \neg D \]

\[ \omega_3 = A \land \neg B \land D \quad \omega_4 = A \land \neg B \land \neg D \]

\[ \omega_5 = \neg A \land B \land D \quad \omega_6 = \neg A \land B \land \neg D \]

\[ \omega_7 = \neg A \land \neg B \land D \quad \omega_8 = \neg A \land \neg B \land \neg D \]

Assignments as constraints over the \(P(\omega_i)\)'s:

\[ p_1 = P(D \mid A) = \frac{P(A \land D)}{P(A)} \]

\[ A \land D = \omega_1 \lor \omega_3, \quad A = \omega_1 \lor \omega_2 \lor \omega_3 \lor \omega_4 \]

hence

\[ p_1 = \frac{[P(\omega_1) + P(\omega_3)]}{[P(\omega_1) + P(\omega_2) + P(\omega_3) + P(\omega_4)]} \]
\[
\begin{align*}
\ p_1 &= \frac{P(\omega_1) + P(\omega_3)}{P(\omega_1) + P(\omega_2) + P(\omega_3) + P(\omega_4)} \\
\ p_2 &= \frac{P(\omega_5) + P(\omega_7)}{P(\omega_1) + P(\omega_2) + P(\omega_3) + P(\omega_4)} \\
\ p_3 &= \frac{P(\omega_1) + P(\omega_5)}{P(\omega_1) + P(\omega_2) + P(\omega_5) + P(\omega_6)} \\
\ p_4 &= \frac{P(\omega_1) + P(\omega_5)}{P(\omega_1) + P(\omega_2) + P(\omega_5) + P(\omega_7)} \\
\ p_5 &= \frac{P(\omega_1) + P(\omega_2)}{P(\omega_1) + P(\omega_2) + P(\omega_3) + P(\omega_4)} \\
\ p_6 &= P(\omega_1) + P(\omega_2) + P(\omega_3) + P(\omega_4) \\
\ 1 &= P(\omega_1) + P(\omega_2) + \ldots + P(\omega_8) \\
\ 0 \leq P(\omega_i) \leq 1 
\end{align*}
\]


The possibilistic model is an alternative approach to uncertainty proposed by Zadeh, and developed by Dubois, Prade and others.

It is more qualitative than quantitative, exploiting plausibility ordering relations on events under incomplete information.
Basic ideas

Assume that we know that:

\[ x \in E \]
\[ E \subseteq A \]
\[ B \cap E \neq \emptyset \]
\[ C \cap E = \emptyset \]

Then we can say that:

- “\( x \in A \)” is certain
- “\( x \in B \)” is possible (but not necessary)
- “\( x \in C \)” is impossible

Inclusions induce necessity, while non-empty intersections only induce possibility.
Basic ideas: Possibility

\[ \Pi(A) = 1 \text{ if } E \cap A \neq \emptyset \]
\[ \Pi(A) = 0 \text{ if } E \cap A = \emptyset \]

Axioms:

\[ \Pi(\emptyset) = 0 \]
\[ \Pi(S) = 1 \]
\[ \Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\} \]

\[ 1 = \Pi(S) = \Pi(A \cup \overline{A}) = \max\{\Pi(A), \Pi(\overline{A})\} \text{: one of } A, \overline{A} \text{ is possible} \]

If \( \Pi(A) = \Pi(\overline{A}) = 1 \): total ignorance
Basic ideas: Necessity

\( A \) is necessary iff \( \overline{A} \) is impossible. \( N(A) = 1 - \Pi(\overline{A}) \)

\[
N(A) = 1 \text{ if } E \subseteq A \\
N(A) = 0 \text{ if } A \subsetneq E
\]

Axioms:

\[
N(\emptyset) = 0 \\
N(S) = 1 \\
N(A \cap B) = \min\{N(A), N(B)\}
\]

\( 0 = N(\emptyset) = N(A \cap \overline{A}) = \min\{N(A), N(\overline{A})\} \): both \( A, \overline{A} \) cannot hold

If \( N(A) = N(\overline{A}) = 0 \): total ignorance
Information: $x \in E$
Possible states: $E$
Impossible states: $S \setminus E$

**Idea:** Refine this simple setting by ranking the states in terms of their plausibility.

Endow $S$ with a partial order relation $\geq_{\Pi}$.
$s_1 \geq_{\Pi} s_2$: $s_1$ is more plausible (less surprising) than $s_2$.

**Equivalent representation:** An ordered partition $\langle E_1, \ldots, E_n \rangle$ of $S$ where $E_1$ are the most plausible states and $E_n$ are the least plausible.
Ordinal setting: examples

A) Uncertain evidence
\( x = \) age of the President

- partial ignorance: he was between 70 and 80 years old when the war started, i.e. \( 70 \leq x \leq 80 \)
- partial ignorance with preferences: he was about 72 when the war started
  \[ "x = 72" >_{\Pi} "x = 71" \sim_{\Pi} "x = 73" >_{\Pi} "x = 70" \sim_{\Pi} "x = 74" >_{\Pi} \ldots \]

B) Uncertain generic information
flying birds \( >_{\Pi} \) non-flying birds
Given $\langle S, \geq_{\Pi} \rangle$ and its corresponding partition $\langle E_1, \ldots, E_n \rangle$, the possibility relation can be extended to subsets $A, B \subseteq S$:

$$A \geq_{\Pi} B \iff \exists u \in A \text{ such that } u \geq_{\Pi} v \forall v \in B \iff \min\{k \mid E_k \cap A \neq \emptyset\} \geq \min\{k \mid E_k \cap B \neq \emptyset\}$$

Dual necessity relation: $A \geq_{N} B \iff \overline{B} \geq_{\Pi} \overline{A}$
The ordered partition \( \langle E_1, \ldots, E_n \rangle \) can be mapped into a plausibility scale on \([0, 1]\)

\[
\pi_x : S \rightarrow [0, 1] \quad \text{possibility distribution}
\]

such that:

- \( \pi_x(u) \geq \pi_x(v) \) iff \( u \geq_{\Pi} v \) (hence: \( \pi_x(u) = \pi_x(v) \) if \( u, v \in E_k \))
- \( \pi_x(u) = 1 \) if \( u \in E_1 \)
- \( \pi_x(u) = 0 \) if \( u \in E_n \)

A possibility distribution represents a state of knowledge, what an agent knows about \( x \)

- \( \pi_x(u) = 0 \): “\( x = u \)” is impossible, excluded
- \( \pi_x(u) = 1 \): “\( x = u \)” is fully plausible, expected
- \( \pi_x(u) > \pi_x(v) \): “\( x = u \)” is more plausible than “\( x = v \)”
A possibility distribution is a fuzzy set over the universe of states.

**Example**

Given the information “the President was about 72 when the war started”, we can take $\pi_x$ as a triangular fuzzy set centered at 72.

Given possibility distributions $\pi, \pi'$ we say that $\pi$ is more specific than $\pi'$ if for every $u \in S$, $\pi(u) \leq \pi'(u)$. $\pi$ involves more commitment (more information) than $\pi'$.

Complete knowledge: $\pi(u_0) = 1$; for $\pi(u) = 0$ for $u \neq u_0$

Total ignorance: $\pi(u) = 1$ for all $u \in S$
A possibility measure on $S$ is a function $\Pi : \mathcal{P}(S) \rightarrow [0, 1]$ such that

$\Pi_1 \quad \Pi(\emptyset) = 0$
$\Pi_2 \quad \Pi(S) = 1$
$\Pi_3 \quad \Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$

A necessity measure on $S$ is a function $N : \mathcal{P}(S) \rightarrow [0, 1]$ such that

$N_1 \quad N(\emptyset) = 0$
$N_2 \quad \Pi(S) = 1$
$N_3 \quad N(A \cap B) = \min\{\Pi(A), \Pi(B)\}$
\(x\): ill-known variable taking values in \(S\)

\(\pi_x: S \rightarrow [0, 1]\) available knowledge about possible values of \(x\)

It induces:

a possibility measure: \(\Pi_x(A) = \max\{\pi_x(u) \mid u \in A\}\)

a necessity measure: \(N_x(A) = 1 - \Pi_x(\overline{A}) = \min\{1 - \pi_x(u) \mid u \notin A\}\)

guaranteed possibility: \(\Delta_x(A) = \min\{\pi_x(u) \mid u \in A\}\)

potential necessity: \(\nabla_x(A) = \max\{1 - \pi_x(u) \mid u \notin A\}\)
Possibilist conditioning – 1

1) Boolean setting: information “$x \in E$” is represented as

$$\pi_x(u) = \begin{cases} 
1 & \text{if } u \in E \\
0 & \text{otherwise}
\end{cases}$$

New information “$x \in A$” (with $A \cap E \neq \emptyset$) updates the representation:

$$\pi_x(u \mid A) = \begin{cases} 
1 & \text{if } u \in E \cap A \\
0 & \text{otherwise}
\end{cases}$$

2) General setting: information represented by $\pi_x : S \to [0, 1]$

New information “$x \in A$” (with $\Pi_x(A) > 0$) updates the representation:

$$\pi_x(u \mid A) = \begin{cases} 
\pi_x(u) & \text{if } u \in A \\
0 & \text{otherwise}
\end{cases}$$
The conditioned distribution may not be normalized:
\[ \Pi_x(S \mid A) = \max_{u \in S} \pi_x(u) = \max_{u \in A} \pi_x(u) = \alpha < 1 \]

**Ordinal normalization:**
\[
\pi'_x(u \mid A) = \begin{cases} 
1 & \text{if } \pi_x(u) = \alpha \\
\pi_x(u) & \text{otherwise}
\end{cases}
\]

\[
\Pi(B \mid A) = \max_{u \in B} \pi'_x(u \mid A) = \begin{cases} 
1 & \text{if } \Pi(B) = \Pi(A \cap B) \\
\Pi(A \cap B) & \text{otherwise}
\end{cases}
\]

**Numerical normalization:**
\[
\pi'_x(u \mid A) = \frac{\pi_x(u \mid A)}{\alpha}
\]

\[
\Pi(B \mid A) = \max_{u \in B} \pi'_x(u \mid A) = \frac{\Pi(A \cap B)}{\Pi(A)}
\]
(Necessity-valued) Possibilistic Logic – 1

Formulae: $\langle \varphi, \alpha \rangle$ where

- $\varphi$ is a classical formula
- $\alpha \in [0, 1]$ is a lower bound on the belief on $\varphi$ in terms of a necessity measure:

$$\langle \varphi, \alpha \rangle \text{ is true iff } N(\varphi) \geq \alpha$$

Examples:

- $\varphi$ is certain $\langle \varphi, 1 \rangle$
- $\varphi$ is $\alpha$-certain $\langle \varphi, \alpha \rangle$
- $\varphi$ is unknown $\langle \varphi, 0 \rangle$
- $\varphi$ is $\beta$-false $\langle \neg \varphi, \beta \rangle$
- $\varphi$ is false $\langle \neg \varphi, 1 \rangle$
(Necessity-valued) Possibilistic Logic – 2

\[ KB = \{ \langle \varphi_i, \alpha_i \rangle \mid i = 1 \ldots n \} \]: possibilistic formulae representing our knowledge

\[ KB^* = \{ \varphi_1, \ldots, \varphi_n \} \]

\[ KB = KB_1 \cup \ldots \cup KB_m \] where \( KB_k = \{ \langle \varphi_{i_k}, \alpha_k \rangle \mid i_k \} \), with priorities \( \alpha_1 > \alpha_2 > \ldots > \alpha_m \)

**Inference:** prioritized extension of classical inference (\( \vdash_{CL} \))

\[ KB \vdash_{PL} \langle \varphi, \alpha_i \rangle \] iff

- \( KB_1^* \cup \ldots \cup KB_i^* \vdash_{CL} \varphi \)
- \( KB_1^* \cup \ldots \cup KB_j^* \not\vdash_{CL} \varphi \) for \( j < i \)
Example of a possibilistic knowledge base – 1

We have an engine with:

- oil pump, fuel pump, motor
- three switches (sw1, sw2, sw3), which can be either ON or OFF

**Evidence:** the three switches are ON, the engine is hot

**Question:** is the engine OK?
Example of a possibilistic knowledge base – 2

\[ KB = \]

- if the fuel pump is clogged, then the motor gets no fuel
- when sw1 is ON, \textit{normally} fuel is pumped properly
- when fuel is pumped properly, the fuel level is \textit{possibly} OK
- when sw2 is ON, \textit{usually} oil is pumped
- when oil is pumped, the oil level is \textit{possibly} OK
- when there is oil and fuel, \textit{usually} the engine works OK
- when there is heat, the engine is \textit{likely} to be not OK
- when there is heat, it is \textit{likely} that there are oil problems
- when fuel is pumped and the speed is low, there are \textit{reasons to believe} that the pump is clogged
- when sw2 is ON, \textit{normally} the speed is low
- when sw2 and sw3 are ON, it is \textit{almost sure} that the speed is not low
- when sw3 is ON, \textit{usually} fuel is OK

Evidence: sw1, sw2, and sw3 are ON and there is heat
Levels of plausibility:

\[ \text{sure} > \text{almost sure} > \text{likely} > \text{usually} \sim \text{normally} > \text{possibly} > \text{reasons} \]

\[ 1 > 0.9 > 0.8 > 0.7 > 0.5 > 0.4 \]
Example of a possibilistic knowledge base – 4

\[
\langle \text{pump\_clog} \rightarrow \neg \text{fuel\_ok}, 1 \rangle \quad \langle \text{sw1} \rightarrow \text{pump\_fuel}, 0.7 \rangle \\
\langle \text{pump\_fuel} \rightarrow \text{fuel\_ok}, 0.5 \rangle \quad \langle \text{sw2} \rightarrow \text{pump\_oil}, 0.7 \rangle \\
\langle \text{pump\_oil} \rightarrow \text{oil\_ok}, 0.5 \rangle \quad \langle \text{fuel\_ok} \land \text{oil\_ok} \rightarrow \text{engine\_ok}, 0.7 \rangle \\
\langle \text{heat} \rightarrow \neg \text{engine\_ok}, 0.8 \rangle \quad \langle \text{heat} \rightarrow \neg \text{oil\_ok}, 0.8 \rangle \\
\langle \text{pump\_fuel} \land \text{low\_speed} \rightarrow \text{pump\_clog}, 0.4 \rangle \quad \langle \text{sw2} \rightarrow \text{low\_speed}, 0.7 \rangle \\
\langle \text{sw2} \land \text{sw3} \rightarrow \neg \text{low\_speed}, 0.9 \rangle \quad \langle \text{sw3} \rightarrow \text{fuel\_ok}, 0.7 \rangle \\
\langle \text{sw1}, 1 \rangle \quad \langle \text{sw2}, 1 \rangle \\
\langle \text{sw3}, 1 \rangle \quad \langle \text{heat}, 1 \rangle 
\]
Example of a possibilistic knowledge base – 5

\[ KB_1 \]
\[ \langle sw1, 1 \rangle \]
\[ \langle sw2, 1 \rangle \]
\[ \langle sw3, 1 \rangle \]
\[ \langle heat, 1 \rangle \]
\[ \langle pump\_clog \rightarrow \neg fuel\_ok, 1 \rangle \]

\[ KB_{0.9} \]
\[ \langle sw2 \land sw3 \rightarrow \neg low\_speed, 0.9 \rangle \]

\[ KB_{0.8} \]
\[ \langle heat \rightarrow \neg engine\_ok, 0.8 \rangle \]
\[ \langle heat \rightarrow \neg oil\_ok, 0.8 \rangle \]
Example of a possibilistic knowledge base – 6

\[
KB_{0.7} = \langle sw1 \rightarrow pump\_fuel, 0.7 \rangle \\
\langle sw2 \rightarrow pump\_oil, 0.7 \rangle \\
\langle fuel\_ok \land oil\_ok \rightarrow engine\_ok, 0.7 \rangle \\
\langle sw2 \rightarrow low\_speed, 0.7 \rangle \\
\langle sw3 \rightarrow fuel\_ok, 0.7 \rangle \\
\]

\[
KB_{0.5} = \langle pump\_fuel \rightarrow fuel\_ok, 0.5 \rangle \\
\langle pump\_oil \rightarrow oil\_ok, 0.5 \rangle \\
\]

\[
KB_{0.4} = \langle pump\_fuel \land low\_speed \rightarrow pump\_clog, 0.4 \rangle \\
\]
$S$: set of classical interpretations
i.e. for each $v \in S$ and each $\varphi$, $v(\varphi) \in \{0, 1\}$
$[\varphi]$: set of classical models of $\varphi$

Belief states:

$$\pi : S \rightarrow [0, 1]$$
$$\pi (v) < \pi (v')$$: $v'$ is more plausible than $v$
Possibilistic Logic: Semantics – 2

Possibilistic satisfaction:

$$\pi \models \langle \varphi, \alpha \rangle \iff N([\varphi] | \pi) \geq \alpha,$$

where

$$N([\varphi] | \pi) = \inf_{v \in S} \max\{1 - \pi(v), \nu(\varphi)\}$$

Observe that:

- $$\pi \models \langle \varphi, \alpha \rangle \iff \pi(v) > 1 - \alpha \Rightarrow v \in [\varphi]$$
- $$\pi \models \langle \varphi, \alpha \rangle \iff \pi \leq \pi_{\langle \varphi, \alpha \rangle}$$

where $$\pi_{\langle \varphi, \alpha \rangle}(v) = \begin{cases} 1 & \text{if } v \in [\varphi] \\ 1 - \alpha & \text{otherwise} \end{cases}$$
$KB = \{ \langle \varphi_i, \alpha_i \rangle \mid i = 1 \ldots n \}$: possibilistic formulae representing our knowledge

Possibilistic semantical consequence: $KB \models_{PL} \langle \psi, \beta \rangle$ iff

for each $\pi : S \to [0, 1]$, $\pi \models KB$ implies $\pi \models \langle \psi, \beta \rangle$
The proof system for PL, denoted as $\vdash_{PL}$, is defined by refutation through resolution.

**Resolution rule:**

$$\frac{\langle \neg p \lor q, \alpha \rangle, \langle p \lor r, \beta \rangle}{\langle q \lor r, \min\{\alpha, \beta\} \rangle}$$

$KB \vdash_{PL} \langle \varphi, \alpha \rangle$ iff there is a proof of $\langle 0, \alpha \rangle$ starting from $KB \cup \{\langle \neg \varphi, 1 \rangle\}$ by applying only the resolution rule.
Theorem (Dubois-Lang-Prade, 1994)

For every possibilistic knowledge base $KB$ and every possibilistic formula $⟨\varphi, \alpha⟩$:

$$KB \vdash^r_{PL} ⟨\varphi, \alpha⟩ \iff KB \models_{PL} ⟨\varphi, \alpha⟩$$
Bibliography