POSSIBILISTIC RULE-BASED SYSTEM: MIN-MAX INFERENCE AND EXPLAINABILITY

GT “Apprentissage et Raisonnement”, GDR IA

Monday 13^{th} December, 2021

Ismaïl Baaj

SCIENCES SORBONNE UNIVERSITÉ

In this presentation, we focus on two objectives:

i) the establishment of meeting points between *Knowledge Representation and Reasoning* (KRR) and *Machine Learning* (ML)

In this presentation, we focus on two objectives:

i) the establishment of meeting points between *Knowledge Representation and Reasoning* (KRR) and *Machine Learning* (ML)

ii) the elaboration of a processing chain for *eXplainable Artificial Intelligence* (XAI) in order to be able to generate AI explanations in natural language and to evaluate them:

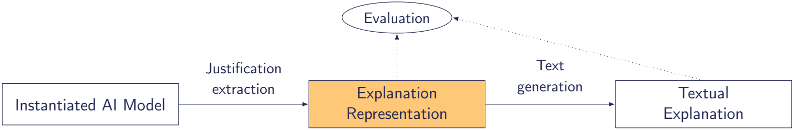

Figure 1: Proposed processing chain to generate and evaluate explanations (Baaj et al. 2019)

Introduction

- Possibility Theory is a well-known framework for the representation of incomplete or imprecise information (*What is possible without being certain at all? What is certain to some extent?*)

- Possibility Theory is a well-known framework for the representation of incomplete or imprecise information (*What is possible without being certain at all? What is certain to some extent?*)

- Dubois and Prade (2020) emphasized the development of possibilistic learning methods that would be *consistent* with if-then rule-based reasoning

SCIENCES SORBONNE UNIVERSITÉ

- Possibility Theory is a well-known framework for the representation of incomplete or imprecise information (*What is possible without being certain at all? What is certain to some extent?*)

- Dubois and Prade (2020) emphasized the development of possibilistic learning methods that would be *consistent* with if-then rule-based reasoning

- Dubois and Prade (2020) highlighted the approach of Farreny and Prade (1989), who proposed a min-max equation system in order to develop the explanatory capabilities of possibilistic rule based systems

Background

- A set of n if-then possibilistic rules $R^1, R^2, \cdots, R^n$
- R^i: "if p_i then q_i" with an uncertainty propagation matrix:

$$\begin{bmatrix} \pi(q_i|p_i) & \pi(q_i|\neg p_i) \\ \pi(\neg q_i|p_i) & \pi(\neg q_i|\neg p_i) \end{bmatrix} = \begin{bmatrix} 1 & s_i \\ r_i & 1 \end{bmatrix}$$

SCIENCES SORBONNE UNIVERSITÉ

- A set of n if-then possibilistic rules $R^1, R^2, \cdots, R^n$
- R^i: "if p_i then q_i" with an uncertainty propagation matrix:

$$\begin{bmatrix} \pi(q_i|p_i) & \pi(q_i|\neg p_i) \\ \pi(\neg q_i|p_i) & \pi(\neg q_i|\neg p_i) \end{bmatrix} = \begin{bmatrix} 1 & s_i \\ r_i & 1 \end{bmatrix}$$

- p_i stands for "$a_i(x) \in P_i$" and q_i for "$b(x) \in Q_i$":

  a_i and b: *attributes* applied to an item x

  P_i and Q_i: *subsets* of the respective *attribute domains* (D_{a_i}, D_b)

- Information about $a_i(x)$ represented by a *possibility distribution* $\pi_{a_i(x)} : D_{a_i} \to [0,1]$ *normalized* i.e. $\exists u \in D_{a_i}$, $\pi_{a_i(x)}(u) = 1$
- Evaluation of p_i: "$a_i(x) \in P_i$":

$$\pi(p_i) = \Pi(P_i) = \sup_{u \in P_i} \pi_{a_i(x)}(u) = \lambda_i$$
$$\pi(\neg p_i) = \Pi(\overline{P_i}) = \sup_{u \in \overline{P_i}} \pi_{a_i(x)}(u) = 1 - n(p_i) = \rho_i$$

where $n(p_i) = N(P_i) = \inf_{u \in \overline{P_i}} (1 - \pi_{a_i(x)}(u)) = 1 - \pi(\neg p_i)$

- Information about $a_i(x)$ represented by a *possibility distribution* $\pi_{a_i(x)} : D_{a_i} \rightarrow [0, 1]$ *normalized* i.e. $\exists u \in D_{a_i}, \pi_{a_i(x)}(u) = 1$
- Evaluation of p_i: "$a_i(x) \in P_i$":

$$\pi(p_i) = \Pi(P_i) = \sup_{u \in P_i} \pi_{a_i(x)}(u) = \lambda_i$$

$$\pi(\neg p_i) = \Pi(\overline{P_i}) = \sup_{u \in \overline{P_i}} \pi_{a_i(x)}(u) = 1 - n(p_i) = \rho_i$$

where $n(p_i) = N(P_i) = \inf_{u \in \overline{P_i}}(1 - \pi_{a_i(x)}(u)) = 1 - \pi(\neg p_i)$

Consequence of the normalization of $\pi_{a_i(x)}$: $\max(\pi(p_i), \pi(\neg p_i)) = 1$

- In case of a compounded premise $p_i = p_{1,i} \wedge \cdots \wedge p_{k,i}$: $\pi(p_i) = \min_{j=1}^{k} \pi(p_{j,i})$ and $\pi(\neg p_i) = \max_{j=1}^{k} \pi(\neg p_{j,i})$

- Uncertainty propagation: $\begin{bmatrix}\pi(q_i)\\ \pi(\neg q_i)\end{bmatrix} = \begin{bmatrix}1 & s_i\\ r_i & 1\end{bmatrix} \Box_{\min}^{\max} \begin{bmatrix}\lambda_i\\ \rho_i\end{bmatrix} = \begin{bmatrix}\alpha_i\\ \beta_i\end{bmatrix}$

$\Box_{\min}^{\max}$: matricial product with min as product and max as addition

- As $\max(\lambda_i, \rho_i) = 1$ we have:

$$\alpha_i = \max(s_i, \lambda_i)$$
$$\beta_i = \max(r_i, \rho_i)$$

- Uncertainty propagation: $\begin{bmatrix}\pi(q_i)\\ \pi(\neg q_i)\end{bmatrix} = \begin{bmatrix}1 & s_i\\ r_i & 1\end{bmatrix} \Box^{\max}_{\min} \begin{bmatrix}\lambda_i\\ \rho_i\end{bmatrix} = \begin{bmatrix}\alpha_i\\ \beta_i\end{bmatrix}$

$\Box^{\max}_{\min}$: matricial product with min as product and max as addition

- As $\max(\lambda_i, \rho_i) = 1$ we have:

$$\alpha_i = \max(s_i, \lambda_i)$$
$$\beta_i = \max(r_i, \rho_i)$$

- Possibility distribution of the attribute b associated to R^i:

$$\pi^{*i}_{b(x)}(u) = \alpha_i \mu_{Q_i}(u) + \beta_i \mu_{\overline{Q_i}}(u) \text{ for any } u \in D_b$$

$\mu_{Q_i}, \mu_{\overline{Q_i}}$: characteristic functions of $Q_i, \overline{Q_i}$

- Possibility distribution of b with n rules:

$$\pi^{*}_{b(x)}(u) = \min(\pi^{*1}_{b(x)}(u), \pi^{*2}_{b(x)}(u), \cdots, \pi^{*n}_{b(x)}(u))$$

- Two sets of possibilistic rules: $R^1, R^2, \cdots, R^n$ and $R'^1, R'^2, \cdots, R'^m$
- Same attribute b used in both the conclusions of the R^i and the premises of the R'^j.

- Two sets of possibilistic rules: $R^1, R^2, \cdots, R^n$ and $R'^1, R'^2, \cdots, R'^m$
- Same attribute b used in both the conclusions of the R^i and the premises of the R'^j.

- Each rule R'_j: "if p'_j then q'_j":

  uncertainty propagation matrix: $\begin{bmatrix} 1 & s'_j \\ r'_j & 1 \end{bmatrix}$

  p'_j stands for "$b(x) \in Q'_j$" and q'_j for "$c(x) \in Q''_j$"

  $Q'_j \subseteq D_b$ and $Q''_j \subseteq D_c$

  $\lambda'_j = \pi(p'_j)$ and $\rho'_j = \pi(\neg p'_j)$

SCIENCES SORBONNE UNIVERSITÉ

Cascade of Farreny and Prade (1989)

- First set of possibilistic rules:

  R^1: if a person likes meeting people, then recommended professions are professor or businessman or lawyer or doctor

  R^2: if a person is fond of creation/inventions, then recommended professions are engineer or public researcher or architect

  - R^3: if a person looks for job security and is fond of intellectual speculation, then recommended professions are professor or public researcher

where $D_{profession} = \{$businessman, lawyer, doctor, professor, researcher, architect, engineer, others$\}$, $s_1 = 1$, $r_1 = 0.3$, $s_2 = 0.2$, $r_2 = 0.4$, $s_3 = 1$, $r_3 = 0.3$.

- Second set of possibilistic rules:

  R'^1: if a person is a professor or a researcher, then her salary is rather low

  R'^2: if a person is an engineer, a lawyer or an architect, her salary is average or high

  R'^3: if a person is a business man or a doctor, then her salary is high

where D_{salary}=$\{$low,average,high$\}$, $s'_1 = 1$, $r'_1 = 0.7$, $s'_2 = 0.8$, $r'_2 = 0.2$, $s'_3 = 0.6$ and $r'_3 = 0.4$.

- Farreny and Prade (1989) proposed an equation system denoted OV = MR ■ IV in order to:
  - describe the output possibility distribution
  - perform a sensitivity analysis

SCIENCES SORBONNE UNIVERSITÉ

- Farreny and Prade (1989) proposed an equation system denoted $OV = MR \blacksquare IV$ in order to:

  describe the output possibility distribution

  perform a sensitivity analysis

- Dubois and Prade (2020) made explicit this equation system for the case of two rules R^1 and R^2:

$$\begin{bmatrix} \Pi(Q_1 \cap Q_2) \\ \Pi(Q_1 \cap \overline{Q_2}) \\ \Pi(\overline{Q_1} \cap Q_2) \\ \Pi(\overline{Q_1} \cap \overline{Q_2}) \end{bmatrix} = \begin{bmatrix} s_1 & 1 & s_2 & 1 \\ s_1 & 1 & 1 & r_2 \\ 1 & r_1 & s_2 & 1 \\ 1 & r_1 & 1 & r_2 \end{bmatrix} \Box^{\min}_{\max} \begin{bmatrix} \lambda_1 \\ \rho_1 \\ \lambda_2 \\ \rho_2 \end{bmatrix} = \begin{bmatrix} \min(\alpha_1, \alpha_2) \\ \min(\alpha_1, \beta_2) \\ \min(\beta_1, \alpha_2) \\ \min(\beta_1, \beta_2) \end{bmatrix}$$

$\Box^{\min}_{\max}$: matricial product with max as product and min as addition

- $Q_1 \cap Q_2, Q_1 \cap \overline{Q_2}, \overline{Q_1} \cap Q_2, \overline{Q_1} \cap \overline{Q_2}$ form a *partition* of D_b

Generalized equation system

- From a possibilistic rule-based system with n rules $R^1, R^2, \cdots, R^n$:

$$O_n = M_n \Box_{\max}^{\min} I_n$$

SCIENCES SORBONNE UNIVERSITÉ

- From a possibilistic rule-based system with n rules $R^1, R^2, \cdots, R^n$:

$$O_n = M_n \Box_{\max}^{\min} I_n$$

- To understand the output vector O_n, we introduce:

  $(E_k^{(n)})_{1 \leq k \leq 2^n}$: an explicit partition of D_b constructed with the sets $Q_1, Q_2, \cdots, Q_n$ used in the conclusions of the rules and their complements

  B_n: a matrix constructed inductively w.r.t the number of rules

- For $i = 1, 2, \cdots, n$, the matrices M_i, I_i, B_i are defined according to:

  $s_1, s_2, \cdots, s_i$ and $r_1, r_2, \cdots, r_i$ for M_i

  $\lambda_1, \lambda_2, \cdots, \lambda_i$ and $\rho_1, \rho_2, \cdots, \rho_i$ for I_i

  $\alpha_1, \alpha_2, \cdots, \alpha_i$ and $\beta_1, \beta_2, \cdots, \beta_i$ for B_i

For each $i = 1, 2, \cdots, n$, the partition $(E_k^{(i)})_{1 \leq k \leq 2^i}$ is defined by the following two conditions:

$E_1^{(1)} = Q_1$ and $E_2^{(1)} = \overline{Q_1}$

and for $i > 1$: $E_k^{(i)} = \begin{cases} E_k^{(i-1)} \cap Q_i & \text{if } 1 \leq k \leq 2^{i-1} \\ E_{k-2^{i-1}}^{(i-1)} \cap \overline{Q_i} & \text{if } 2^{i-1} < k \leq 2^i \end{cases}$

- Respective size of M_i, I_i and B_i: $(2^i, 2i)$, $(2i, 1)$ and $(2^i, i)$
- $i = 1$, we take $M_1 = \begin{bmatrix} s_1 & 1 \\ 1 & r_1 \end{bmatrix}$, $I_1 = \begin{bmatrix} \lambda_1 \\ \rho_1 \end{bmatrix}$, $B_1 = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix}$

- Respective size of M_i, I_i and B_i: $(2^i, 2i)$, $(2i, 1)$ and $(2^i, i)$
- $i = 1$, we take $M_1 = \begin{bmatrix} s_1 & 1 \\ 1 & r_1 \end{bmatrix}$, $I_1 = \begin{bmatrix} \lambda_1 \\ \rho_1 \end{bmatrix}$, $B_1 = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix}$
- $i > 1$, we define $M_i = \left[\begin{array}{c|c} M_{i-1} & S_i \\ \hline M_{i-1} & R_i \end{array}\right]$, $I_i = \left[\begin{array}{c} I_{i-1} \\ \hline \lambda_i \\ \rho_i \end{array}\right]$, $B_i = \left[\begin{array}{c|c} & \alpha_i \\ & \alpha_i \\ B_{i-1} & \vdots \\ & \alpha_i \\ \hline & \beta_i \\ & \beta_i \\ B_{i-1} & \vdots \\ & \beta_i \end{array}\right]$

where $S_i = \begin{bmatrix} s_i & 1 \\ s_i & 1 \\ \vdots & \vdots \\ s_i & 1 \end{bmatrix}$ and $R_i = \begin{bmatrix} 1 & r_i \\ 1 & r_i \\ \vdots & \vdots \\ 1 & r_i \end{bmatrix}$ of size $(2^{i-1}, 2)$

- $E_k^{(i)}$ is linked to the row $L_k = (\gamma_1, \gamma_2, \cdots, \gamma_i)$ of B_i with $\gamma \in \{\alpha, \beta\}$ by:

$$E_k^{(i)} = T_1 \cap T_2 \cdots \cap T_i \text{ with } T_j = \begin{cases} Q_j & \text{if } \gamma_j = \alpha_j \\ \overline{Q_j} & \text{if } \gamma_j = \beta_j \end{cases}$$

Generalized equation system – *Matrix* $\boxdot_{\min} B_i$

- For any $i = 1, 2, \cdots, n$, we set:

$$\boxdot_{\min} B_i = [o_k^{(i)}]_{1 \leq k \leq 2^i}$$

$\boxdot_{\min}$: the minimum of the coefficients of each row in a matrix

- For any $k \in \{1, 2, \cdots, 2^i\}$, we deduce:

$$o_k^{(i)} = \begin{cases} \min(o_k^{(i-1)}, \alpha_i) & \text{if } 1 \leq k \leq 2^{i-1} \\ \min(o_{k-2^{i-1}}^{(i-1)}, \beta_i) & \text{if } 2^{i-1} < k \leq 2^i \end{cases}$$

- Finally, we obtain:

Theorem

$$M_i \Box_{\max}^{\min} I_i = \boxdot_{\min} B_i$$

A possibilistic rule-based system composed of $n = 3$ rules

- The sets of the partition $(E_k^{(3)})_{1 \leq k \leq 8}$ are the following: $Q_1 \cap Q_2 \cap Q_3$, $\overline{Q}_1 \cap Q_2 \cap Q_3$, $Q_1 \cap \overline{Q}_2 \cap Q_3$, $\overline{Q}_1 \cap \overline{Q}_2 \cap Q_3$, $Q_1 \cap Q_2 \cap \overline{Q}_3$, $\overline{Q}_1 \cap Q_2 \cap \overline{Q}_3$, $Q_1 \cap \overline{Q}_2 \cap \overline{Q}_3$ and $\overline{Q}_1 \cap \overline{Q}_2 \cap \overline{Q}_3$

- We check the Theorem by direct calculation:

$O_3 = M_3 \Box_{\max}^{\min} I_3$

$$= \begin{bmatrix} s_1 & 1 & s_2 & 1 & s_3 & 1 \\ 1 & r_1 & s_2 & 1 & s_3 & 1 \\ s_1 & 1 & 1 & r_2 & s_3 & 1 \\ 1 & r_1 & 1 & r_2 & s_3 & 1 \\ s_1 & 1 & s_2 & 1 & 1 & r_3 \\ 1 & r_1 & s_2 & 1 & 1 & r_3 \\ s_1 & 1 & 1 & r_2 & 1 & r_3 \\ 1 & r_1 & 1 & r_2 & 1 & r_3 \end{bmatrix} \Box_{\max}^{\min} \begin{bmatrix} \lambda_1 \\ \rho_1 \\ \lambda_2 \\ \rho_2 \\ \lambda_3 \\ \rho_3 \end{bmatrix} = \begin{bmatrix} \min(\alpha_1, \alpha_2, \alpha_3) \\ \min(\beta_1, \alpha_2, \alpha_3) \\ \min(\alpha_1, \beta_2, \alpha_3) \\ \min(\beta_1, \beta_2, \alpha_3) \\ \min(\alpha_1, \alpha_2, \beta_3) \\ \min(\beta_1, \alpha_2, \beta_3) \\ \min(\alpha_1, \beta_2, \beta_3) \\ \min(\beta_1, \beta_2, \beta_3) \end{bmatrix} = \boxdot_{\min} B_3$$

Equation system properties

- Using the coefficients of $O_i = \boxdot_{\min} B_i$ and the characteristic functions $\mu_{E_1^{(i)}}, \mu_{E_2^{(i)}}, \cdots, \mu_{E_{2^i}^{(i)}}$ of the sets $E_1^{(i)}, E_2^{(i)}, \cdots, E_{2^i}^{(i)}$:

Theorem

The output possibility distribution $\pi^*_{b(x),i}$ associated to the first i rules is:

$$\pi^*_{b(x),i} = \sum_{1 \leq k \leq 2^i} o_k^{(i)} \mu_{E_k^{(i)}}$$

- Using the coefficients of $O_i = \boxdot_{\min} B_i$ and the characteristic functions $\mu_{E_1^{(i)}}, \mu_{E_2^{(i)}}, \cdots, \mu_{E_{2^i}^{(i)}}$ of the sets $E_1^{(i)}, E_2^{(i)}, \cdots, E_{2^i}^{(i)}$:

Theorem

The output possibility distribution $\pi^*_{b(x),i}$ associated to the first i rules is:

$$\pi^*_{b(x),i} = \sum_{1 \leq k \leq 2^i} o_k^{(i)} \mu_{E_k^{(i)}}$$

- Consequence: $\forall u \in D_b$, $\exists k_0$ unique s.t $u \in E_{k_0}^{(i)}$ and $\pi^*_{b(x),i}(u) = o_{k_0}^{(i)}$
- $\pi^*_{b(x),i}$ is normalized iff: $\exists k \in \{1, 2, \cdots, 2^i\}$ s.t $E_k^{(i)} \neq \emptyset$ and $o_k^{(i)} = 1$

Example (continued) : a possibilistic rule-based system composed of $n = 3$ rules

- The characteristic functions of the partition $(E_k^{(3)})_{1 \leq k \leq 8}$ are $\mu_{Q_1 \cap Q_2 \cap Q_3}$, $\mu_{\overline{Q_1} \cap Q_2 \cap Q_3}$, $\mu_{Q_1 \cap \overline{Q_2} \cap Q_3}$, $\mu_{\overline{Q_1} \cap \overline{Q_2} \cap Q_3}$, $\mu_{Q_1 \cap Q_2 \cap \overline{Q_3}}$, $\mu_{\overline{Q_1} \cap Q_2 \cap \overline{Q_3}}$, $\mu_{Q_1 \cap \overline{Q_2} \cap \overline{Q_3}}$ and $\mu_{\overline{Q_1} \cap \overline{Q_2} \cap \overline{Q_3}}$
- The output possibility distribution is:

$$\begin{aligned}
\pi^*_{b(x),3} &= \min(\pi^{*1}_{b(x)}, \pi^{*2}_{b(x)}, \pi^{*3}_{b(x)}) \\
&= \min(\alpha_1, \alpha_2, \alpha_3)\mu_{Q_1 \cap Q_2 \cap Q_3} + \min(\beta_1, \alpha_2, \alpha_3)\mu_{\overline{Q_1} \cap Q_2 \cap Q_3} \\
&\quad + \min(\alpha_1, \beta_2, \alpha_3)\mu_{Q_1 \cap \overline{Q_2} \cap Q_3} + \min(\beta_1, \beta_2, \alpha_3)\mu_{\overline{Q_1} \cap \overline{Q_2} \cap Q_3} \\
&\quad + \min(\alpha_1, \alpha_2, \beta_3)\mu_{Q_1 \cap Q_2 \cap \overline{Q_3}} + \min(\beta_1, \alpha_2, \beta_3)\mu_{\overline{Q_1} \cap Q_2 \cap \overline{Q_3}} \\
&\quad + \min(\alpha_1, \beta_2, \beta_3)\mu_{Q_1 \cap \overline{Q_2} \cap \overline{Q_3}} + \min(\beta_1, \beta_2, \beta_3)\mu_{\overline{Q_1} \cap \overline{Q_2} \cap \overline{Q_3}}
\end{aligned}$$

- J contains the indexes of the non-empty sets of the partition:

$$J = \{k \in \{1, 2, \cdots, 2^i\} \mid E_k^{(i)} \neq \emptyset\} \text{ and } \omega = \text{card}(J)$$

Arrange the elements of J as a strictly increasing sequence:
$1 \leq k_1 < k_2 < \cdots < k_\omega \leq 2^i$

We have $\omega \leq \min(d, 2^i)$ where $d = \text{card}(D_b)$

$[\Pi(E_k^{(i)})]_{k \in J} = [o_k^{(i)}]_{k \in J}$

- Let $\mathcal{O}_i$, $\mathcal{M}_i$ and $\mathcal{B}_i$ be the matrices obtained from O_i, M_i and B_i respectively, by deleting each row whose index is not in J

- First set of possibilistic rules of the cascade of Farreny and Prade ($n = 3$):

> R^1: if a person likes meeting people, then recommended professions are professor or business man or lawyer or doctor
> R^2: if a person is fond of creation/inventions, then recommended professions are engineer or public researcher or architect
> R^3: if a person looks for job security and is fond of intellectual speculation,then recommended professions are professor or public researcher

where $D_{profession} = \{$business man,lawyer, doctor, professor, researcher, architect, engineer, others$\}$, $s_1 = 1$, $r_1 = 0.3$, $s_2 = 0.2$, $r_2 = 0.4$, $s_3 = 1$, $r_3 = 0.3$

- Partition of $D_{profession}$: $E^{(3)}_{k_1} = \{\text{researcher}\}$, $E^{(3)}_{k_2} = \{\text{professor}\}$, $E^{(3)}_{k_3} = \{\text{engineer}, \text{architect}\}$, $E^{(3)}_{k_4} = \{\text{business man}, \text{lawyer}, \text{doctor}\}$ and $E^{(3)}_{k_5} = \{\text{others}\}$
- Equation system:

$$\begin{bmatrix} \Pi(E^{(3)}_{k_1}) \\ \Pi(E^{(3)}_{k_2}) \\ \Pi(E^{(3)}_{k_3}) \\ \Pi(E^{(3)}_{k_4}) \\ \Pi(E^{(3)}_{k_5}) \end{bmatrix} = \begin{bmatrix} 1 & r_1 & s_2 & 1 & s_3 & 1 \\ s_1 & 1 & 1 & r_2 & s_3 & 1 \\ 1 & r_1 & s_2 & 1 & 1 & r_3 \\ s_1 & 1 & 1 & r_2 & 1 & r_3 \\ 1 & r_1 & 1 & r_2 & 1 & r_3 \end{bmatrix} \Box^{\min}_{\max} \begin{bmatrix} \lambda_1 \\ \rho_1 \\ \lambda_2 \\ \rho_2 \\ \lambda_3 \\ \rho_3 \end{bmatrix}$$

The vector $\mathcal{O}_3$ and the matrix $\mathcal{M}_3$ have <u>five rows</u> (while O_3 and M_3 have eight rows).

Equation system properties

- Let: $\varepsilon(T) = \begin{cases} 1 & \text{si } T \neq \emptyset \\ 0 & \text{if } T = \emptyset \end{cases}$
- To any matrix $A = [a_{ij}]$, we associate $A^\circ = [1 - a_{ij}]$. $(A^\circ)^\circ = A$.

- Let: $\varepsilon(T) = \begin{cases} 1 & \text{si } T \neq \emptyset \\ 0 & \text{if } T = \emptyset \end{cases}$
- To any matrix $A = [a_{ij}]$, we associate $A^\circ = [1 - a_{ij}]$. $(A^\circ)^\circ = A$.
- For $Q \subseteq D_b$, we have $Q = \bigcup_{1 \leq j \leq \omega} E^{(i)}_{k_j} \cap Q$.

Equation system properties

- Let: $\varepsilon(T) = \begin{cases} 1 & \text{si } T \neq \emptyset \\ 0 & \text{if } T = \emptyset \end{cases}$
- To any matrix $A = [a_{ij}]$, we associate $A^\circ = [1 - a_{ij}]$. $(A^\circ)^\circ = A$.
- For $Q \subseteq D_b$, we have $Q = \bigcup_{1 \leq j \leq \omega} E_{k_j}^{(i)} \cap Q$.

The possibility measure is:

$$\Pi^*(Q) = \max_{u \in Q} \pi^*_{b(x)}(u) = \nabla_Q \Box_{\min}^{\max} \mathcal{O}_i$$

where $\nabla_Q = \begin{bmatrix} \varepsilon(E_{k_1}^{(i)} \cap Q) & \varepsilon(E_{k_2}^{(i)} \cap Q) & \cdots & \varepsilon(E_{k_\omega}^{(i)} \cap Q) \end{bmatrix}$

- Let: $\varepsilon(T) = \begin{cases} 1 & \text{si } T \neq \emptyset \\ 0 & \text{if } T = \emptyset \end{cases}$
- To any matrix $A = [a_{ij}]$, we associate $A^\circ = [1 - a_{ij}]$. $(A^\circ)^\circ = A$.
- For $Q \subseteq D_b$, we have $Q = \bigcup_{1 \leq j \leq \omega} E^{(i)}_{k_j} \cap Q$.

The possibility measure is:

$$\Pi^*(Q) = \max_{u \in Q} \pi^*_{b(x)}(u) = \nabla_Q \Box^{\max}_{\min} \mathcal{O}_i$$

where $\nabla_Q = \left[\varepsilon(E^{(i)}_{k_1} \cap Q) \quad \varepsilon(E^{(i)}_{k_2} \cap Q) \quad \cdots \quad \varepsilon(E^{(i)}_{k_\omega} \cap Q)\right]$

- As $\Pi^*(\overline{Q}) = \nabla_{\overline{Q}} \Box^{\max}_{\min} \mathcal{O}_i$, the necessity measure is then:

$$N^*(Q) = 1 - \Pi^*(\overline{Q}) = (\Pi^*(\overline{Q}))^\circ$$

By the correspondences between the operators $\Box^{\min}_{\max}$ and $\Box^{\max}_{\min}$ we obtain:

$$N^*(Q) = (\nabla_{\overline{Q}} \Box^{\max}_{\min} \mathcal{O}_i)^\circ = \nabla_{\overline{Q}}{}^\circ \Box^{\min}_{\max} \mathcal{O}_i{}^\circ$$

SCIENCES SORBONNE UNIVERSITÉ

First set of possibilistic rules of the cascade of Farreny and Prade:

- Partition of $D_{profession}$: $E^{(3)}_{k_1} = \{\text{researcher}\}$, $E^{(3)}_{k_2} = \{\text{professor}\}$, $E^{(3)}_{k_3} = \{\text{engineer}, \text{architect}\}$, $E^{(3)}_{k_4} = \{\text{business man}, \text{lawyer}, \text{doctor}\}$ and $E^{(3)}_{k_5} = \{\text{others}\}$
- Equation system with $\lambda_1 = 1, \rho_1 = 0.5, \lambda_2 = 0.2, \rho_2 = 1, \lambda_3 = 1, \rho_3 = 0.6$:

$$\begin{bmatrix}\Pi(E^{(3)}_{k_1})\\ \Pi(E^{(3)}_{k_2})\\ \Pi(E^{(3)}_{k_3})\\ \Pi(E^{(3)}_{k_4})\\ \Pi(E^{(3)}_{k_5})\end{bmatrix} = \begin{bmatrix}1 & r_1 & s_2 & 1 & s_3 & 1\\ s_1 & 1 & 1 & r_2 & s_3 & 1\\ 1 & r_1 & s_2 & 1 & 1 & r_3\\ s_1 & 1 & 1 & r_2 & 1 & r_3\\ 1 & r_1 & 1 & r_2 & 1 & r_3\end{bmatrix} \Box^{\min}_{\max} \begin{bmatrix}\lambda_1\\ \rho_1\\ \lambda_2\\ \rho_2\\ \lambda_3\\ \rho_3\end{bmatrix} = \begin{bmatrix}0.2\\ 1\\ 0.2\\ 0.6\\ 0.5\end{bmatrix}$$

Let $Q = \{\text{professor}, \text{researcher}\}$. The possibility measure of Q is:

$$\Pi^*(Q) = \nabla_Q \Box^{\max}_{\min} \mathcal{O}_3 = 1$$

where $\nabla_Q = \begin{bmatrix}\varepsilon(E^{(i)}_{k_1} \cap Q) & \varepsilon(E^{(i)}_{k_2} \cap Q) & \cdots & \varepsilon(E^{(i)}_{k_\omega} \cap Q)\end{bmatrix} = \begin{bmatrix}1 & 1 & 0 & 0 & 0\end{bmatrix}$

The necessity measure of Q is 0.4.

Cascade and applications

- Two equation systems:

$$\mathcal{O}_n = \mathcal{M}_n \Box_{\max}^{\min} I_n \text{ for } R^1, R^2, \cdots, R^n$$
$$\mathcal{O}'_m = \mathcal{M}'_m \Box_{\max}^{\min} I'_m \text{ for } R'^1, R'^2, \cdots, R'^m$$

- The input vector I'_m is linked to the output vector $\mathcal{O}_n$ by:

$$I'_m = \nabla \Box_{min}^{max} \mathcal{O}_n \quad \text{where} \quad \nabla = \begin{bmatrix} \nabla_{Q'_1} \\ \nabla_{\overline{Q'_1}} \\ \vdots \\ \nabla_{Q'_m} \\ \nabla_{\overline{Q'_m}} \end{bmatrix}$$

- The output vector O'_m is deduced from the first system:

$$\begin{aligned} \mathcal{O}'_m &= \mathcal{M}'_m \Box_{\max}^{\min} I'_m \\ &= \mathcal{M}'_m \Box_{\max}^{\min} (\nabla \Box_{\min}^{\max} \mathcal{O}_n) \\ &= \mathcal{M}'_m \Box_{\max}^{\min} (\nabla \Box_{\min}^{\max} (\mathcal{M}_n \Box_{\max}^{\min} I_n)) \end{aligned}$$

SCIENCES SORBONNE UNIVERSITÉ

- First set of possibilistic rules:

  R^1: if a person likes meeting people, then recommended professions are professor or business man or lawyer or doctor

  R^2: if a person is fond of creation/inventions, then recommended professions are engineer or public researcher or architect

  R^3: if a person looks for job security and is fond of intellectual speculation,then recommended professions are professor or public researcher

where $D_{profession} = \{$business man,lawyer, doctor, professor, researcher, architect, engineer, others$\}$, $s_1 = 1$, $r_1 = 0.3$, $s_2 = 0.2$, $r_2 = 0.4$, $s_3 = 1$, $r_3 = 0.3$

- Partition of $D_{profession}$: $E^{(3)}_{k_1} = \{\text{researcher}\}$, $E^{(3)}_{k_2} = \{\text{professor}\}$,

$E^{(3)}_{k_3} = \{\text{engineer}, \text{architect}\}$, $E^{(3)}_{k_4} = \{\text{business man}, \text{lawyer}, \text{doctor}\}$ and $E^{(3)}_{k_5} = \{\text{others}\}$

- Equation system with $\lambda_1 = 1, \rho_1 = 0.5, \lambda_2 = 0.2, \rho_2 = 1, \lambda_3 = 1, \rho_3 = 0.6$:

$$\begin{bmatrix} \Pi(E^{(3)}_{k_1}) \\ \Pi(E^{(3)}_{k_2}) \\ \Pi(E^{(3)}_{k_3}) \\ \Pi(E^{(3)}_{k_4}) \\ \Pi(E^{(3)}_{k_5}) \end{bmatrix} = \begin{bmatrix} 1 & r_1 & s_2 & 1 & s_3 & 1 \\ s_1 & 1 & 1 & r_2 & s_3 & 1 \\ 1 & r_1 & s_2 & 1 & 1 & r_3 \\ s_1 & 1 & 1 & r_2 & 1 & r_3 \\ 1 & r_1 & 1 & r_2 & 1 & r_3 \end{bmatrix} \Box^{\min}_{\max} \begin{bmatrix} \lambda_1 \\ \rho_1 \\ \lambda_2 \\ \rho_2 \\ \lambda_3 \\ \rho_3 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 1 \\ 0.2 \\ 0.6 \\ 0.5 \end{bmatrix}$$

- Second set of possibilistic rules:

  R'^1: if a person is a professor or a researcher, then her salary is rather low
  R'^2: if a person is an engineer, a lawyer or an architect, her salary is average or high
  R'^3: if a person is a business man or a doctor, then her salary is high

where D_{salary}={low,average,high}, $s'_1 = 1$, $r'_1 = 0.7$, $s'_2 = 0.8$, $r'_2 = 0.2$, $s'_3 = 0.6$ and $r'_3 = 0.4$

- Second set of possibilistic rules:

  R'^1: if a person is a professor or a researcher, then her salary is rather low
  R'^2: if a person is an engineer, a lawyer or an architect, her salary is average or high
  R'^3: if a person is a business man or a doctor, then her salary is high

where D_{salary}={low,average,high}, $s'_1 = 1$, $r'_1 = 0.7$, $s'_2 = 0.8$, $r'_2 = 0.2$, $s'_3 = 0.6$ and $r'_3 = 0.4$

- Partition of D_{salary}: $E'^{(3)}_{k_1} = \{\text{high}\}$, $E'^{(3)}_{k_2} = \{\text{average}\}$ and $E'^{(3)}_{k_3} = \{\text{low}\}$
- $$I'_m = \nabla \Box^{max}_{min} \mathcal{O}_3 = \begin{bmatrix} \nabla_{Q'_1} \\ \nabla_{\overline{Q'_1}} \\ \nabla_{Q'_2} \\ \nabla_{\overline{Q'_2}} \\ \nabla_{Q'_3} \\ \nabla_{\overline{Q'_3}} \end{bmatrix} \Box^{max}_{min} \mathcal{O}_3 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \Box^{max}_{min} \begin{bmatrix} 0.2 \\ 1 \\ 0.2 \\ 0.6 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.6 \\ 0.6 \\ 1 \\ 0.6 \\ 1 \end{bmatrix}$$
- Equation system:

$$\begin{bmatrix} \Pi(E'^{(3)}_{k_1}) \\ \Pi(E'^{(3)}_{k_2}) \\ \Pi(E'^{(3)}_{k_3}) \end{bmatrix} = \begin{bmatrix} 1 & r'_1 & s'_2 & 1 & s'_3 & 1 \\ 1 & r'_1 & s'_2 & 1 & 1 & r'_3 \\ s'_1 & 1 & 1 & r'_2 & 1 & r'_3 \end{bmatrix} \Box^{min}_{max} \begin{bmatrix} \lambda'_1 \\ \rho'_1 \\ \lambda'_2 \\ \rho'_2 \\ \lambda'_3 \\ \rho'_3 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.7 \\ 1 \end{bmatrix}$$

- (Reminder) The output vector O'_m is deduced from the first system:

$$\mathcal{O}'_m = \mathcal{M}'_m \Box^{\min}_{\max} (\nabla \Box^{\max}_{\min} (\mathcal{M}_n \Box^{\min}_{\max} I_n))$$

- With the matrices $I_n{}^\circ$, $\mathcal{M}_n{}^\circ$ and $\mathcal{M}'_m{}^\circ$, we can express the equations involved in the cascade using only the operator $(A \Box^{\max}_{\min} B)^\circ$:

$$\mathcal{O}_n = (\mathcal{M}_n{}^\circ \Box^{\max}_{\min} I_n{}^\circ)^\circ$$
$$I'_m{}^\circ = (\nabla \Box^{\max}_{\min} \mathcal{O}_n)^\circ$$
$$\mathcal{O}'_m = (\mathcal{M}'_m{}^\circ \Box^{\max}_{\min} I'_m{}^\circ)^\circ$$

- We have:

$$\mathcal{O}'_m = (\mathcal{M}'_m{}^\circ \Box^{\max}_{\min} (\nabla \Box^{\max}_{\min} (\mathcal{M}_n{}^\circ \Box^{\max}_{\min} I_n{}^\circ)^\circ)^\circ)^\circ$$

- The cascade construction is represented by a min-max neural network:

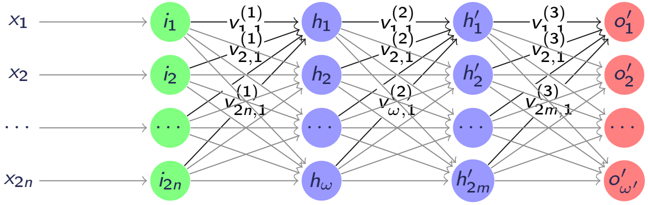

Figure 2: Min-max neural network architecture

- For a neuron x linked by t edges $v_1, v_2, \cdots, v_t$ to t ancestors, whose output values are $y_1, y_2, \cdots, y_t$ we compute:
  - its input value: $1 - \max_{1 \le j \le t} \min(v_j, y_j) = \min_{1 \le j \le t} \max(1 - v_j, 1 - y_j)$
  - its output value: $f(x) = x$

In my PhD thesis:

- We study the *existence of a minimal input vector* for $\pi^*_{b(x)}(u) = 1$
- We define an algorithm to rebuild the equation system *when we remove a rule*. It outputs the equation system associated to the remaining subset of rules.
  Therefore, it enables us to obtain all the equation subsystems of an initial equation system

Perspectives:

- Sensitivity analysis and explainability, as suggested by Farreny and Prade (1989)
- Coherence of the rule base: conditions on the parameters of the rules and the input vector
- Possibilistic learning: a min-max gradient descent method may be developed for our neural network
- The learning of the parameters of the rules s_i, r_i may be done with the help of the algorithms for solving systems of fuzzy relational equations (Sanchez 1977, Peeva 2013)

Explainability: justifying inference results

- Representation of the information given by the possibility and necessity degrees of a premise p of a rule "if p then q":

Notation

For a premise p, the triplet (p, sem, d) *denotes either* $(p, P, \pi(p))$ *or* $(p, C, n(p))$, *where* $sem \in \{P, C\}$ *(P for possible, C for certain) is the semantics attached to the degree* $d \in \{\pi(p), n(p)\}$

- Representation of the information given by the possibility and necessity degrees of a premise p of a rule "if p then q":

Notation

For a premise p, the triplet (p, sem, d) denotes either $(p, P, \pi(p))$ or $(p, C, n(p))$, where $sem \in \{P, C\}$ (P for possible, C for certain) is the semantics attached to the degree $d \in \{\pi(p), n(p)\}$

- Possibility degree of an output attribute value u computed by:

$$\pi^*_{b(x)}(u) = \min(\gamma_1, \gamma_2, \dots, \gamma_n), \quad (1)$$

where $\gamma_i = \pi^{*i}_{b(x)}(u) = \max(t_i, \theta_i)$ with $(t_i, \theta_i) = \begin{cases} (s_i, \lambda_i) & \text{if } \gamma_i = \alpha_i \\ (r_i, \rho_i) & \text{if } \gamma_i = \beta_i \end{cases}$

- Triplets according to the $\gamma_1, \gamma_2, \dots, \gamma_n$ appearing in the relation (1):

$$(p_i, sem_i, d_i) = \begin{cases} (p_i, \mathrm{P}, \lambda_i) & \text{if} \quad \gamma_i = \alpha_i \\ (p_i, \mathrm{C}, 1 - \rho_i) & \text{if} \quad \gamma_i = \beta_i \end{cases} \quad (2)$$

Example

- Blood sugar control system for a patient with type 1 diabetes:

	activity (act)	current-bloodsugar (cbs)	future-bloodsugar (fbs)
R^1	dinner, drink-coffee, lunch	medium, high	high
R^2	long-sleep, sport, walking	low, medium	low
R^3	alcohol-consumption, breakfast	low, medium	low, medium

Table 1: Rule base for the control of the blood sugar level.

We have: $D_{\text{act}} = \{$*alcohol-consumption*, *breakfast*, *dinner*, *drink-coffee*, *long-sleep*, *lunch*, *sport*, *walking*$\}$ and $D_{\text{cbs}} = D_{\text{fbs}} = \{$*low*, *medium*, *high*$\}$.

- Blood sugar control system for a patient with type 1 diabetes:

	activity (act)	current-bloodsugar (cbs)	future-bloodsugar (fbs)
R^1	dinner, drink-coffee, lunch	medium, high	high
R^2	long-sleep, sport, walking	low, medium	low
R^3	alcohol-consumption, breakfast	low, medium	low, medium

Table 1: Rule base for the control of the blood sugar level.

We have: $D_{\text{act}} = \{alcohol\text{-}consumption, breakfast, dinner, drink\text{-}coffee, long\text{-}sleep, lunch, sport, walking\}$ and $D_{\text{cbs}} = D_{\text{fbs}} = \{low, medium, high\}$.

- By the relation (1), the possibility degree of the output values low, medium and high are:

$\pi_{\text{fbs}(x)}(\text{low}) = \min(\gamma_1^l, \gamma_2^l, \gamma_3^l)$, where $\gamma_1^l = \beta_1 = \max(r_1, \rho_1)$, $\gamma_2^l = \alpha_2 = \max(s_2, \lambda_2)$, $\gamma_3^l = \alpha_3 = \max(s_3, \lambda_3)$.

$\pi_{\text{fbs}(x)}(\text{medium}) = \min(\gamma_1^m, \gamma_2^m, \gamma_3^m)$, where $\gamma_1^m = \beta_1$, $\gamma_2^m = \beta_2 = \max(r_2, \rho_2)$, $\gamma_3^m = \alpha_3$.

$\pi_{\text{fbs}(x)}(\text{high}) = \min(\gamma_1^h, \gamma_2^h, \gamma_3^h)$, where $\gamma_1^h = \alpha_1 = \max(s_1, \lambda_1)$, $\gamma_2^h = \beta_2$, $\gamma_3^h = \beta_3 = \max(r_3, \rho_3)$.

- Blood sugar control system for a patient with type 1 diabetes:

	activity (act)	current-bloodsugar (cbs)	future-bloodsugar (fbs)
R^1	dinner, drink-coffee, lunch	medium, high	high
R^2	long-sleep, sport, walking	low, medium	low
R^3	alcohol-consumption, breakfast	low, medium	low, medium

Table 1: Rule base for the control of the blood sugar level.

We have: $D_{\text{act}} = \{$*alcohol-consumption*, *breakfast*, *dinner*, *drink-coffee*, *long-sleep*, *lunch*, *sport*, *walking*$\}$ and $D_{\text{cbs}} = D_{\text{fbs}} = \{low, medium, high\}$.

- By the relation (1), the possibility degree of the output values low, medium and high are:

$\pi_{\text{fbs}(x)}(\text{low}) = \min(\gamma_1^l, \gamma_2^l, \gamma_3^l)$, where $\gamma_1^l = \beta_1 = \max(r_1, \rho_1)$, $\gamma_2^l = \alpha_2 = \max(s_2, \lambda_2)$, $\gamma_3^l = \alpha_3 = \max(s_3, \lambda_3)$.
$\pi_{\text{fbs}(x)}(\text{medium}) = \min(\gamma_1^m, \gamma_2^m, \gamma_3^m)$, where $\gamma_1^m = \beta_1$, $\gamma_2^m = \beta_2 = \max(r_2, \rho_2)$, $\gamma_3^m = \alpha_3$.
$\pi_{\text{fbs}(x)}(\text{high}) = \min(\gamma_1^h, \gamma_2^h, \gamma_3^h)$, where $\gamma_1^h = \alpha_1 = \max(s_1, \lambda_1)$, $\gamma_2^h = \beta_2$, $\gamma_3^h = \beta_3 = \max(r_3, \rho_3)$.

- Using Notation 1, the following triplets are set (relation 2):

for $\pi_{\text{fbs}(x)}(\text{low})$: $(p_1, \text{C}, 1 - \rho_1), (p_2, \text{P}, \lambda_2)$, $(p_3, \text{P}, \lambda_3)$.
for $\pi_{\text{fbs}(x)}(\text{medium})$: $(p_1, \text{C}, 1 - \rho_1), (p_2, \text{C}, 1 - \rho_2)$, $(p_3, \text{P}, \lambda_3)$.
for $\pi_{\text{fbs}(x)}(\text{high})$: $(p_1, \text{P}, \lambda_1), (p_2, \text{C}, 1 - \rho_2)$, $(p_3, \text{C}, 1 - \rho_3)$.

Farreny and Prade's approach focuses on two explanatory purposes for an output attribute value $u \in D_b$:

(i) How to get $\pi^*_{b(x)}(u)$ strictly greater or lower than a given $\tau \in [0, 1]$?

Farreny and Prade's approach focuses on two explanatory purposes for an output attribute value $u \in D_b$:

(i) How to get $\pi^*_{b(x)}(u)$ strictly greater or lower than a given $\tau \in [0,1]$?
Reminder: the parameters of the rules are set:
$\pi^*_{b(x)}(u)$ ranges between $\omega = \min(t_1, t_2, \ldots, t_n)$ and 1

- for $\pi^*_{b(x)}(u) > \tau$: $\forall i \in \{j \in \{1,2,\ldots,n\} \mid t_j \leq \tau\}$ we have $\theta_i > \tau$
- for $\pi^*_{b(x)}(u) < \tau$ with $\omega < \tau \leq 1$: $\exists i \in \{j \in \{1,2,\ldots,n\} \mid t_j < \tau\}$ s.t. $\theta_i < \tau$

Farreny and Prade's approach focuses on two explanatory purposes for an output attribute value $u \in D_b$:

(i) How to get $\pi^*_{b(x)}(u)$ strictly greater or lower than a given $\tau \in [0,1]$?
Reminder: the parameters of the rules are set:
$\pi^*_{b(x)}(u)$ ranges between $\omega = \min(t_1, t_2, \ldots, t_n)$ and 1

- for $\pi^*_{b(x)}(u) > \tau$: $\forall i \in \{j \in \{1,2,\ldots,n\} \mid t_j \leq \tau\}$ we have $\theta_i > \tau$
- for $\pi^*_{b(x)}(u) < \tau$ with $\omega < \tau \leq 1$: $\exists i \in \{j \in \{1,2,\ldots,n\} \mid t_j < \tau\}$ s.t. $\theta_i < \tau$

Example

For our blood sugar control system with $s_1 = 1, r_2 = r_3 = 0$, we have $\pi^*_{fbs(x)}(high) > 0.5$ iff $\rho_2 > 0.5$ and $\rho_3 > 0.5$

Farreny and Prade's approach focuses on two explanatory purposes for an output attribute value $u \in D_b$:

(i) How to get $\pi^*_{b(x)}(u)$ strictly greater or lower than a given $\tau \in [0, 1]$?
Reminder: the parameters of the rules are set:
$\pi^*_{b(x)}(u)$ ranges between $\omega = \min(t_1, t_2, \dots, t_n)$ and 1

- for $\pi^*_{b(x)}(u) > \tau$: $\forall i \in \{j \in \{1, 2, \dots, n\} \mid t_j \leq \tau\}$ we have $\theta_i > \tau$
- for $\pi^*_{b(x)}(u) < \tau$ with $\omega < \tau \leq 1$: $\exists i \in \{j \in \{1, 2, \dots, n\} \mid t_j < \tau\}$ s.t. $\theta_i < \tau$

Example

For our blood sugar control system with $s_1 = 1, r_2 = r_3 = 0$, we have $\pi^*_{fbs(x)}(high) > 0.5$ iff $\rho_2 > 0.5$ and $\rho_3 > 0.5$

(ii) What are the degrees of the premises justifying $\pi^*_{b(x)}(u) = \tau$?

- Let two sets compare the parameters $t_1, t_2, \cdots t_n$ of the rules to the degrees $\theta_1, \theta_2, \cdots, \theta_n$:

$J^P = \{i \in \{1, 2, \cdots, n\} \mid t_i \leq \theta_i\}$ and $J^R = \{i \in \{1, 2, \cdots, n\} \mid t_i \geq \theta_i\}$

We have $\{1, 2, \cdots, n\} = J^P \cup J^R$ but J^P or J^R may be empty

- Let two sets compare the parameters $t_1, t_2, \cdots t_n$ of the rules to the degrees $\theta_1, \theta_2, \cdots, \theta_n$:

$J^P = \{i \in \{1, 2, \cdots, n\} \mid t_i \leq \theta_i\}$ and $J^R = \{i \in \{1, 2, \cdots, n\} \mid t_i \geq \theta_i\}$

We have $\{1, 2, \cdots, n\} = J^P \cup J^R$ but J^P or J^R may be empty

- With the convention $\min_\emptyset = 1$, we take:
  - $c_\theta = \min_{i \in J^P} \theta_i$: the *lowest possibility degree justifiable by premises* (if $J^P \neq \emptyset$)
  - $c_t = \min_{i \in J^R} t_i$: the *lowest possibility degree justifiable by the parameters of the rules* (if $J^R \neq \emptyset$)

- Let two sets compare the parameters $t_1, t_2, \cdots t_n$ of the rules to the degrees $\theta_1, \theta_2, \cdots, \theta_n$:

$$J^P = \{i \in \{1, 2, \cdots, n\} \mid t_i \leq \theta_i\} \text{ and } J^R = \{i \in \{1, 2, \cdots, n\} \mid t_i \geq \theta_i\}$$

We have $\{1, 2, \cdots, n\} = J^P \cup J^R$ but J^P or J^R may be empty

- With the convention $\min_\emptyset = 1$, we take:
  - $c_\theta = \min_{i \in J^P} \theta_i$: the *lowest possibility degree justifiable by premises* (if $J^P \neq \emptyset$)
  - $c_t = \min_{i \in J^R} t_i$: the *lowest possibility degree justifiable by the parameters of the rules* (if $J^R \neq \emptyset$)
- By using the properties of the min function, we establish:

Proposition

$$\tau = \min(c_\theta, c_t)$$

- *When we can't explain by degrees of premises*: if $J^P = \emptyset$

As the degrees $\theta_1, \theta_2, \cdots, \theta_n$ of the premises are computed using the possibility distributions of the input attributes, we may have $J^P = \emptyset$. In that case, $c_\theta = 1$, $J^R = \{1, 2, \cdots, n\}$ and:

$$\pi^*_{b(x)}(u) = c_t = \min(t_1, t_2, \cdots, t_n)$$

SCIENCES SORBONNE UNIVERSITÉ

- Blood sugar control system for a patient with type 1 diabetes:

	activity (act)	current-bloodsugar (cbs)	future-bloodsugar (fbs)
R^1	dinner, drink-coffee, lunch	medium, high	high
R^2	long-sleep, sport, walking	low, medium	low
R^3	alcohol-consumption, breakfast	low, medium	low, medium

Table 2: Rule base for the control of the blood sugar level.

Inputs: the patient wants to drink a coffee and his current blood sugar level is medium:

$$\pi_{act(x)}(drink\text{-}coffee) = 1, \quad \pi_{cbs(x)}(medium) = 1 \quad \text{and} \quad \pi_{cbs(x)}(low) = 0.3$$

while the other elements of the domains of the input attributes have a possibility degree equal to zero. We have $s_1 = 1, s_2 = 0.7, s_3 = 1, r_1 = r_2 = r_3 = 0.$ and $\lambda_1 = 1$, $\rho_1 = 0.3$, $\lambda_2 = 0$, $\rho_2 = 1$, $\lambda_3 = 0$ and $\rho_3 = 1$. The obtained output possibility distribution is:

$$\pi^*_{fbs(x)}(low) = 0.3, \quad \pi^*_{fbs(x)}(medium) = 0.3 \quad \text{and} \quad \pi^*_{fbs(x)}(high) = 1$$

- Blood sugar control system for a patient with type 1 diabetes:

	activity (act)	current-bloodsugar (cbs)	future-bloodsugar (fbs)
R^1	dinner, drink-coffee, lunch	medium, high	high
R^2	long-sleep, sport, walking	low, medium	low
R^3	alcohol-consumption, breakfast	low, medium	low, medium

Table 2: Rule base for the control of the blood sugar level.

Inputs: the patient wants to drink a coffee and his current blood sugar level is medium:

$$\pi_{act(x)}(drink\text{-}coffee) = 1, \quad \pi_{cbs(x)}(medium) = 1 \quad \text{and} \quad \pi_{cbs(x)}(low) = 0.3$$

while the other elements of the domains of the input attributes have a possibility degree equal to zero. We have $s_1 = 1, s_2 = 0.7, s_3 = 1, r_1 = r_2 = r_3 = 0.$ and $\lambda_1 = 1$, $\rho_1 = 0.3$, $\lambda_2 = 0$, $\rho_2 = 1$, $\lambda_3 = 0$ and $\rho_3 = 1$. The obtained output possibility distribution is:

$$\pi^*_{fbs(x)}(low) = 0.3, \quad \pi^*_{fbs(x)}(medium) = 0.3 \quad \text{and} \quad \pi^*_{fbs(x)}(high) = 1$$

We form the following sets for each output attribute value and deduce their respective c_θ, c_t:

- for *low*: $J_l^P = \{1\}$ and $J_l^R = \{2, 3\}$: $c_\theta^l = 0.3 = \pi_{\mathbf{fbs(x)}}(\text{low})$ and $c_t^l = 0.7$
- for *medium*: $J_m^P = \{1, 2\}$ and $J_m^R = \{3\}$: $c_\theta^m = 0.3 = \pi_{\mathbf{fbs(x)}}(\text{medium})$ and $c_t^m = 1$
- for *high*: $J_h^P = \{1, 2, 3\}$ and $J_h^R = \{1\}$: $c_\theta^h = c_t^h = 1 = \pi_{\mathbf{fbs(x)}}(\text{high})$

Extraction of premises justifying $\pi^*_{b(x)}(u) = \tau$

To explain the inference results of our possibilistic rule-based system, we introduce a threshold $\eta > 0$:

Definition

If a possibility (resp. necessity) degree is higher than the threshold η, it intuitively means that the information it models is relevantly possible (resp. certain)

To explain the inference results of our possibilistic rule-based system, we introduce a threshold $\eta > 0$:

Definition

If a possibility (resp. necessity) degree is higher than the threshold η, it intuitively means that the information it models is relevantly possible (resp. certain)

For a given output value $u \in D_b$, we extract the rule premises justifying the possibility degree $\pi^*_{b(x)}(u) = \tau$ by the following formula:

$$J_{b(x)}(u) = \begin{cases} \{(p_i, \text{sem}_i, d_i) \mid i \in J^P \text{ and } \theta_i = \tau\} & \text{if } \tau \geq \eta \\ \{(p_i, \text{sem}_i, d_i) \mid i \in \{1, 2, \ldots, n\} \text{ and } \gamma_i < \eta\} & \text{if } \tau < \eta \end{cases}$$

Notes: if $\tau \geq \eta$, we rely on J^P (which may be empty) and the condition $\tau = c_\theta$
Otherwise, if $\tau < \eta$ it always exists at least a premise justifying $\pi^*_{b(x)}(u)$

SCIENCES SORBONNE UNIVERSITÉ

- Blood sugar control system for a patient with type 1 diabetes:

	activity (act)	current-bloodsugar (cbs)	future-bloodsugar (fbs)
R^1	dinner, drink-coffee, lunch	medium, high	high
R^2	long-sleep, sport, walking	low, medium	low
R^3	alcohol-consumption, breakfast	low, medium	low, medium

Table 3: Rule base for the control of the blood sugar level.

We have $s_1 = 1, s_2 = 0.7, s_3 = 1, r_1 = r_2 = r_3 = 0.$ and $\lambda_1 = 1$, $\rho_1 = 0.3$, $\lambda_2 = 0$, $\rho_2 = 1$, $\lambda_3 = 0$ and $\rho_3 = 1$. The obtained output possibility distribution is:

$$\pi^*_{fbs(x)}(low) = 0.3, \quad \pi^*_{fbs(x)}(medium) = 0.3 \quad \text{and} \quad \pi^*_{fbs(x)}(high) = 1$$

Let us take $\eta = 0.1$. We obtain for each output attribute value:

- $J_{\mathrm{fbs(x)}}(\mathrm{low}) = J_{\mathrm{fbs(x)}}(\mathrm{medium}) = \{(p_1, \mathrm{C}, 0.7)\}$
- $J_{\mathrm{fbs(x)}}(\mathrm{high}) = \{(p_1, \mathrm{P}, 1), (p_2, \mathrm{C}, 0), (p_3, \mathrm{C}, 0)\}$

If instead of $r_1 = 0$, we take $r_1 > 0.3$, then for $u = low$, the corresponding set J^P is empty: no justification in terms of premises could be given in that case

Justification and unexpectedness of $\pi^*_{b(x)}(u)$

Purpose : for an output attribute value $u \in D_b$, apply reduction functions ($\mathscr{R}_\pi, \mathscr{R}_n, \mathscr{C}_\pi, \mathscr{C}_n$) to the selected premises in $J_{b(x)}(u)$ in order to **form explanations of** $\pi^*_{b(x)}(u)$:

SCIENCES SORBONNE UNIVERSITÉ

Purpose : for an output attribute value $u \in D_b$, apply reduction functions ($\mathscr{R}_\pi, \mathscr{R}_n, \mathscr{C}_\pi, \mathscr{C}_n$) to the selected premises in $J_{b(x)}(u)$ in order to **form explanations of** $\pi^*_{b(x)}(u)$:

- Using $\mathscr{R}_\pi, \mathscr{R}_n$: **the justification of** $\pi^*_{b(x)}(u)$: *A set of possibilistic expressions that are sufficient to justify "b(x) is u at* $\pi^*_{b(x)}(u)$*"*

Purpose : for an output attribute value $u \in D_b$, apply reduction functions $(\mathscr{R}_\pi, \mathscr{R}_n, \mathscr{C}_\pi, \mathscr{C}_n)$ to the selected premises in $J_{b(x)}(u)$ in order to **form explanations of** $\pi^*_{b(x)}(u)$:

- Using $\mathscr{R}_\pi, \mathscr{R}_n$: **the justification of** $\pi^*_{b(x)}(u)$: *A set of possibilistic expressions that are sufficient to justify "b(x) is u at* $\pi^*_{b(x)}(u)$*"*
- Using $\mathscr{C}_\pi, \mathscr{C}_n$: **the unexpectedness of** $\pi^*_{b(x)}(u)$: *A set of possible or certain possibilistic expressions, which may appear to be incompatible with* $\pi^*_{b(x)}(u)$ *while not being involved in its determination*
  - In Simplicity Theory (Dessalles 2008), the unexpectedness aims at capturing exactly what people consider *surprising* in a given situation
  - An unexpectedness X let us formulate statements such as:

    "*even if X, b(x) is u at* $\pi^*_{b(x)}(u)$"
  - It is in the same vein as the "even-if-because" of (Darwiche 2020)

Justification and unexpectedness of $\pi^*_{b(x)}(u)$

Preliminaries : definition of two proposition reduction functions $\mathscr{P}_\pi$, $\mathscr{P}_n$

Justification and unexpectedness of $\pi^*_{b(x)}(u)$

Preliminaries : definition of two proposition reduction functions $\mathscr{P}_\pi$, $\mathscr{P}_n$

- Let a be an attribute with a normalized possibility distribution $\pi_{a(x)}$ on D_a and a proposition p of the form "$a(x) \in P$", where $P \subseteq D_a$.
- We introduce the following two subsets of D_a:

$$(P)_\pi = \{v \in P \mid \pi(v) = \Pi(P)\} \text{ related to a proposition } p_\pi,$$
$$(P)_n = P \cup \{v \in \overline{P} \mid 1 - \pi(v) > N(P)\} \text{ related to } p_n$$

We have $\overline{(P)_n} = \{v \in \overline{P} \mid 1 - \pi(v) = N(P)\}$, $\overline{(P)_n} = (\overline{P})_\pi$, and $\overline{(P)_\pi} = (\overline{P})_n$. Therefore:

$$(P)_n = \overline{(\overline{P})_\pi} \text{ and } (P)_\pi = \overline{(\overline{P})_n}$$

Example

Let us take the possibility distribution π on $D_{cbs} = \{low, medium, high\}$ defined by:

$$\pi(low) = 0.3, \quad \pi(medium) = 1, \quad \pi(high) = 0.$$

Given $P = \{medium, high\}$, we have $(P)_\pi = \{medium\}$ and $(P)_n = P$.
For $P' = \{low\}$, we have $(P')_\pi = P'$ and $(P')_n = \{low, high\}$

Preliminaries : definition of two proposition reduction functions $\mathscr{P}_\pi$, $\mathscr{P}_n$

- Let a be an attribute with a normalized possibility distribution $\pi_{a(x)}$ on D_a and a proposition p of the form "$a(x) \in P$", where $P \subseteq D_a$.
- We introduce the following two subsets of D_a:

$$(P)_\pi = \{v \in P \mid \pi(v) = \Pi(P)\} \text{ related to a proposition } p_\pi,$$
$$(P)_n = P \cup \{v \in \overline{P} \mid 1 - \pi(v) > N(P)\} \text{ related to } p_n$$

We have $\overline{(P)_n} = \{v \in \overline{P} \mid 1 - \pi(v) = N(P)\}$, $\overline{(P)_n} = (\overline{P})_\pi$, and $\overline{(P)_\pi} = (\overline{P})_n$. Therefore:

$$(P)_n = \overline{(\overline{P})_\pi} \text{ and } (P)_\pi = \overline{(\overline{P})_n}$$

Preliminaries : definition of two proposition reduction functions $\mathscr{P}_\pi$, $\mathscr{P}_n$

- Let a be an attribute with a normalized possibility distribution $\pi_{a(x)}$ on D_a and a proposition p of the form "$a(x) \in P$", where $P \subseteq D_a$.
- We introduce the following two subsets of D_a:

$$(P)_\pi = \{v \in P \mid \pi(v) = \Pi(P)\} \text{ related to a proposition } p_\pi,$$
$$(P)_n = P \cup \{v \in \overline{P} \mid 1 - \pi(v) > N(P)\} \text{ related to } p_n$$

We have $\overline{(P)_n} = \{v \in \overline{P} \mid 1 - \pi(v) = N(P)\}$, $\overline{(P)_n} = (\overline{P})_\pi$, and $\overline{(P)_\pi} = (\overline{P})_n$. Therefore:

$$(P)_n = \overline{(\overline{P})_\pi} \text{ and } (P)_\pi = \overline{(\overline{P})_n}$$

- Definitions: $\mathscr{P}_\pi$ reduces P if $\pi(p) \geq \eta$ and $\mathscr{P}_n$ reduces $\overline{P}$ if $n(p) \geq \eta$:

$$\mathscr{P}_\pi(p) = \begin{cases} p_\pi & \text{if } \pi(p) \geq \eta \\ p & \text{if } \pi(p) < \eta \end{cases} \text{ and } \mathscr{P}_n(p) = \begin{cases} p_n & \text{if } n(p) \geq \eta \\ p & \text{if } n(p) < \eta \end{cases}$$

- $\pi(\mathscr{P}_\pi(p)) = \pi(p)$ and $n(\mathscr{P}_n(p)) = n(p)$

- Let $p = p_1 \wedge p_2 \wedge \cdots \wedge p_k$ be a compounded premise, where p_j for $j = 1, 2, \cdots, k$, is a proposition of the form "$a_j(x) \in P_j$" with $P_j \subseteq D_{a_j}$

- $\mathscr{R}_\pi$ returns the structure responsible for $\pi(p)$:

$$\mathscr{R}_\pi(p) = \begin{cases} \bigwedge_{j=1}^{k} \mathscr{P}_\pi(p_j) & \text{if } \pi(p) \geq \eta \\ \bigwedge_{p_j \in \{p_s \mid \pi(p_s) < \eta \text{ for } s=1,\cdots,k\}} p_j & \text{if } \pi(p) < \eta \end{cases}$$

- $\mathscr{R}_n$ returns the structure responsible for $n(p)$:

$$\mathscr{R}_n(p) = \begin{cases} \bigwedge_{j=1}^{k} \mathscr{P}_n(p_j) & \text{if } n(p) \geq \eta \\ \bigwedge_{p_j \in \{p_s \mid n(p_s) < \eta \text{ for } s=1,\cdots,k\}} p_j & \text{if } n(p) < \eta \end{cases}$$

- $\pi(\mathscr{R}_\pi(p)) = \pi(p)$ and $n(\mathscr{R}_n(p)) = n(p)$

Premise reductions functions

- Let $p = p_1 \wedge p_2 \wedge \cdots \wedge p_k$ be a compounded premise, where p_j for $j = 1, 2, \cdots, k$, is a proposition of the form "$a_j(x) \in P_j$" with $P_j \subseteq D_{a_j}$

- When $\pi(p) < \eta$ and $A_p^\pi = \{p_j \mid \pi(p_j) \geq \eta \text{ for } j = 1, \cdots, k\} \neq \emptyset$, $\mathscr{C}_\pi$ returns a conjunction of propositions, which is *not* involved in the determination of $\pi(p)$:

$$\mathscr{C}_\pi(p) = \bigwedge_{p_j \in A_p^\pi} \mathscr{P}_\pi(p_j)$$

- When $n(p) < \eta$ and $A_p^n = \{p_j \mid n(p_j) \geq \eta \text{ for } j = 1, \cdots, k\} \neq \emptyset$, $\mathscr{C}_n$ returns a conjunction of propositions, which is *not* involved in the determination of $n(p)$:

$$\mathscr{C}_n(p) = \bigwedge_{p_j \in A_p^n} \mathscr{P}_n(p_j)$$

- If $\pi(p) < \eta$, (resp. $n(p) < \eta$), each proposition p_j composing p, is either used in $\mathscr{R}_\pi(p)$ or in $\mathscr{C}_\pi(p)$ (resp. $\mathscr{R}_n(p)$ or in $\mathscr{C}_n(p)$), according to $\pi(p_j)$ (resp. $n(p_j)$)

SCIENCES SORBONNE UNIVERSITÉ

- Blood sugar control system for a patient with type 1 diabetes:

	activity (act)	current-bloodsugar (cbs)	future-bloodsugar (fbs)
R^1	dinner, drink-coffee, lunch	medium, high	high
R^2	long-sleep, sport, walking	low, medium	low
R^3	alcohol-consumption, breakfast	low, medium	low, medium

Table 4: Rule base for the control of the blood sugar level.

We have $s_1 = 1, s_2 = 0.7, s_3 = 1, r_1 = r_2 = r_3 = 0.$ and $\lambda_1 = 1$, $\rho_1 = 0.3$, $\lambda_2 = 0$, $\rho_2 = 1$, $\lambda_3 = 0$ and $\rho_3 = 1$. Let us take $\eta = 0.1$.

- For the proposition "act(x) $\in$ {dinner, drink-coffee, lunch}" of p_1 of R^1:

  $\mathscr{P}_\pi$ reduces it to "act(x) $\in$ {drink-coffee}"

  $\mathscr{P}_n$ keeps it as is

- For the premise p_2:"act(x) $\in$ {long-sleep, sport, walking} and cbs(x) $\in$ {low, medium}" of R^2:

  $\mathscr{R}_\pi$ returns the proposition "act(x) $\in$ {long-sleep, sport, walking}"

  $\mathscr{C}_\pi$ returns "cbs(x) $\in$ {medium}"

Apply reduction functions to the premise of a triplet

Reminder: $J_{b(x)}(u)$ is composed of triplets (p, sem, d)

- To apply the reduction functions $\mathscr{R}_\pi$ and $\mathscr{R}_n$ to the premise p of a triplet (p, sem, d), we introduce the function $\mathscr{S}_\mathscr{R}$:

$$\mathscr{S}_\mathscr{R}(p, sem, d) = \begin{cases} (\mathscr{R}_\pi(p), sem, d) \text{ if } sem = \mathrm{P} \\ (\mathscr{R}_n(p), sem, d) \text{ if } sem = \mathrm{C} \end{cases}$$

- Similarly, to apply $\mathscr{C}_\pi$ and $\mathscr{C}_n$, we introduce the function $\mathscr{S}_\mathscr{C}$:

$$\mathscr{S}_\mathscr{C}(p, sem, d) = \begin{cases} (\mathscr{C}_\pi(p), sem, \pi(\mathscr{C}_\pi(p))) \text{ if } sem = \mathrm{P}, d < \eta \text{ and } A_p^\pi \neq \emptyset \\ (\mathscr{C}_n(p), sem, n(\mathscr{C}_n(p))) \text{ if } sem = \mathrm{C}, d < \eta \text{ and } A_p^n \neq \emptyset \end{cases}$$

SCIENCES SORBONNE UNIVERSITÉ

- **The justification of** $\pi^*_{b(x)}(u)$: A set of possibilistic expressions that are sufficient to justify "b(x) is u at $\pi^*_{b(x)}(u)$":

$$\text{Justification}_{b(x)}(u) = \{\mathscr{S}_{\mathscr{R}}(p, sem, d) \mid (p, sem, d) \in J_{b(x)}(u)\}$$

Justification and unexpectedness of $\pi^*_{b(x)}(u)$

- **The justification of** $\pi^*_{b(x)}(u)$: A set of possibilistic expressions that are sufficient to justify "b(x) is u at $\pi^*_{b(x)}(u)$":

$$\text{Justification}_{b(x)}(u) = \{\mathscr{S}_{\mathscr{R}}(p, sem, d) \mid (p, sem, d) \in J_{b(x)}(u)\}$$

- **The unexpectedness of** $\pi^*_{b(x)}(u)$: a set of possible or certain possibilistic expressions, which may appear to be incompatible with $\pi^*_{b(x)}(u)$ while not being involved in its determination:

$$\text{Unexpectedness}_{b(x)}(u) = \{\mathscr{S}_{\mathscr{C}}(p, sem, d) \mid (p, sem, d) \in J_{b(x)}(u)\}$$

Representation of Explanations

- Purpose: represent graphically explanations of possibilistic inference decisions by **conceptual graphs** (Chein & Mugnier 2008)

- Purpose: represent graphically explanations of possibilistic inference decisions by **conceptual graphs** (Chein & Mugnier 2008)

- Representation of a *possibilistic expression* of an explanation (justification, unexpectedness) by a possibilistic conceptual graph :

Definition

A possibilistic conceptual graph (PCG) is a basic conceptual graph (BG) $G = (C, R, E, l)$, where C is the concept nodes set, R the relation nodes set, E is the multi-edges set and the label function l is extended by allowing a degree and a semantics in the label of any concept node $c \in C$:

$$l(c) = (\text{type}(c) : \text{marker}(c) | sem_c, d_c),$$

where $sem_c \in \{\text{P}, \text{C}\}$

- The definition of a *star BG* i.e., a BG restricted to a relation node and its neighbors, is naturally extended as a star PCG

From an explanation (justification, unexpectedness), we form an ontology (called: a *vocabulary* in the conceptual graph framework) to build $m+1$ *possibilistic conceptual graphs* (m : number of possibilistic expressions in an explanation) and a *basic conceptual graph*:

From an explanation (justification, unexpectedness), we form an ontology (called: a *vocabulary* in the conceptual graph framework) to build $m+1$ *possibilistic conceptual graphs* (m : number of possibilistic expressions in an explanation) and a *basic conceptual graph*:

- D : PCG representing the observed phenomenon : the possibility degree $\pi^*_{b(x)}(u)$
- $N_1, \cdots, N_m$: star PCGs representing the m extracted possibilistic expressions of an explanation
- R (*root*) : star BG structuring the explanation by representing the link (*isJustifiedBy* or *evenIf*) between D and $N_1, \cdots, N_m$

The representation is a **nested conceptual graph** G by its associated tree $Tree(G)$ where the graphs $D, N_1, \cdots, N_m$ are nested in R:

Definition

$Tree(G) = (V_T, U_T, l_T)$ is given by:
• $V_T = \{R, D, N_1, N_2, \cdots, N_m\}$ is the set of nodes,
• $U_T = \{(R, D), (R, N_1), (R, N_2), \cdots, (R, N_m)\}$ is the set of edges and the node R is the root of $Tree(G)$,
• the labels of the edges are given by $l_T(R, D) = (R, c_0, D)$ and $l_T(R, N_i) = (R, c_i, N_i)$ for $i = 1, 2, \cdots, m$.

Example (continued)

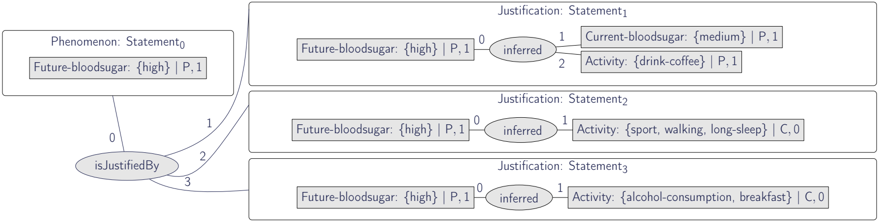

Figure 3: Representation of an explanation

A natural language explanation could be: "*It is possible that the patient's blood sugar level will become high. In fact, his activity is drinking coffee and his current blood sugar level is medium. In addition, it is assessed as not certain that he chose sport, walking, sleeping, eating breakfast or drinking alcohol as an activity.*"

Combining Justification and Unexpectedness

In my PhD thesis, the framework is extended to represent an explanation that is a combination of a justification and an unexpectedness:

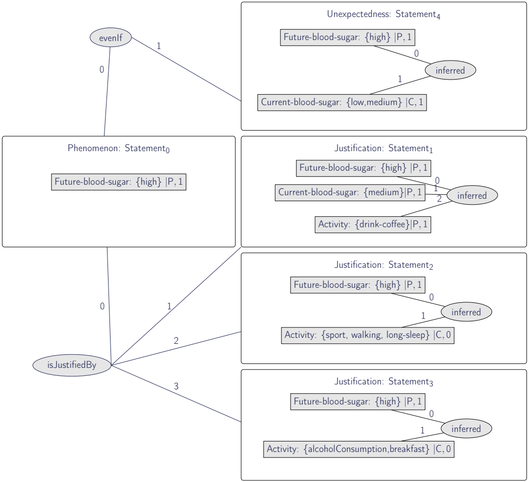

Figure 4: Representation: combination of the justification of $\pi^*_{fbs(x)}(high)$ and its unexpectedness

Conclusion

Conclusion

- Canonical construction for the matrices governing the min-max equation system of Farreny and Prade (1989)
- Formulas for $\pi^*_{b(x)}$ and its possibility and necessity measures
- Representation of the cascade by a min-max neural network

- Canonical construction for the matrices governing the min-max equation system of Farreny and Prade (1989)
- Formulas for $\pi^*_{b(x)}$ and its possibility and necessity measures
- Representation of the cascade by a min-max neural network
- Necessary and sufficient condition for justifying by rule premises the possibility degree $\pi^*_{b(x)}(u)$
- Two explanations of $\pi^*_{b(x)}(u)$: the justification and the unexpectedness
- Representation of explanations of possibilistic inference decisions by nested conceptual graphs, which may be used by natural language generation systems

References

References I

Dubois, D., Prade, H. (2020, September). From Possibilistic Rule-Based Systems to Machine Learning-A Discussion Paper. In International Conference on Scalable Uncertainty Management (pp. 35-51). Springer, Cham.

Farreny, H., Prade, H. (1992). Positive and negative explanations of uncertain reasoning in the framework of possibility theory. In Fuzzy logic for the management of uncertainty (pp. 319-333).

Darwiche, Adnan, and Auguste Hirth. "On the Reasons Behind Decisions." ECAI 2020. IOS Press, 2020. 712-720.

Dimulescu, Adrian, and Jean-Louis Dessalles. "Understanding narrative interest: Some evidence on the role of unexpectedness." (2009).

Chein, M., Mugnier, M.L.: Graph-based knowledge representation: computational foundations of conceptual graphs. Springer Science Business Media (2008).