

Inférence min-max pour un système à base de règles possibilistes

Ismaïl Baaj, Jean-Philippe Poli, Wassila Ouerdane, Nicolas Maudet

▶ **To cite this version:**

Ismaïl Baaj, Jean-Philippe Poli, Wassila Ouerdane, Nicolas Maudet. Inférence min-max pour un système à base de règles possibilistes. Rencontres francophones sur la logique floue et ses applications, Université de la Sorbonne, Oct 2021, Paris, France. pp.233-240. cea-03402616

HAL Id: cea-03402616
https://cea.hal.science/cea-03402616

Submitted on 25 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Inférence min-max pour un système à base de règles possibilistes

Ismaïl Baaj[1,2] Jean-philippe Poli[1] Wassila Ouerdane[3] Nicolas Maudet[2]

[1] Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France.
{ismail.baaj, jean-philippe.poli}@cea.fr

[2] LIP6, Sorbonne Université, Paris, France.
{ismail.baaj, nicolas.maudet}@lip6.fr

[3] MICS, CentraleSupélec, Université Paris-Saclay, Gif sur Yvette, France.
wassila.ouerdane@centralesupelec.fr

Résumé :

Dans cet article, nous explorons le mécanisme d'inférence min-max d'un système de n règles possibilistes. Nous établissons une formule additive pour la distribution des possibilités de sortie obtenue par l'inférence. À partir de ce résultat, nous calculons les mesures de possibilité et de nécessité correspondantes. Nous donnons des conditions nécessaires et suffisantes pour obtenir la normalisation de la distribution des possibilités de sortie. Comme application de nos résultats, nous étudions le cas d'une cascade de deux ensembles de règles possibilistes et établissons une relation liant leurs systèmes d'équations respectifs. Enfin, nous associons à une telle cascade un réseau de neurones min-max.

Mots-clés :

Théorie des possibilités, système à base de règles, réseau de neurones

Abstract:

In this paper, we give a canonical construction for the matrices governing the min-max equation system of Farreny and Prade, which is associated to a rule-based system composed of n parallel possibilistic rules. From this construction and with the help of a partition of the output attribute domain, we establish an additive formula for the output possibility distribution and deduce the corresponding possibility and necessity measures. We give necessary and sufficient conditions for the normalization of the output possibility distribution. In addition, we tackle the case of a cascade and establish for it an input-output relation between the two min-max equation systems. Finally, we represent this cascade construction by a min-max neural network.

Keywords:

Possibility theory, rule-based system, neural network

1 Introduction

La théorie des possibilités permet de représenter des informations incomplètes ou imprécises à l'aide de deux mesures duales appelées *possibilité* et *nécessité* [2, 3]. Ces mesures permettent de distinguer ce qui est possible sans être du tout certain et ce qui est certain dans une moindre mesure. Récemment, Dubois et Prade ont affirmé que la théorie des possibilités offre un cadre intéressant pour le développement d'interfaces cohérentes entre l'apprentissage et le raisonnement [4].

Initialement introduite par Zadeh [7], la théorie des possibilités a été considérablement développée par Dubois et Prade. Cela a conduit à l'étude de systèmes à base de règles possibilistes [6]. Les capacités explicatives de ces systèmes ont été développées par Farreny et Prade [5]. Ces auteurs ont introduit un système d'équations min-max pour effectuer une analyse de sensibilité. Ils ont illustré leurs travaux à l'aide d'un exemple d'un système composé de trois règles parallèles. Ces travaux ont récemment été repris par Dubois et Prade pour développer des méthodes d'apprentissage possibilistes cohérentes avec le raisonnement basé sur les règles [4]. Dans ce but, les auteurs de [4] soulignent que l'on peut tirer profit du système d'équations min-max de Farreny et Prade [5] et l'importance d'un tel système d'équations pour une cascade i.e., un système utilisant consécutivement deux ensembles de règles possibilistes, les règles du second ensemble étant chaînées avec celles du premier ensemble.

Dans cet article, nous étudions un certain nombre de questions ou de problèmes mentionnés dans [4] et [5]. Dans la section 2, nous commençons par rappeler le mécanisme d'inférence d'un système à base de règles possibilistes et le système d'équations de Farreny et Prade [5]. Dans le cas de n règles possibilistes, nous donnons (section 3) une construction canonique des

matrices gouvernant ce système d'équations. Ceci nous permet d'établir une formule additive pour la distribution des possibilités de sortie (section 4). La distribution des possibilités de sortie doit être normalisée pour assurer la cohérence des règles et traiter une cascade. En utilisant notre formule additive, nous donnons une condition nécessaire et suffisante pour que la distribution des possibilités de sortie soit normalisée (section 4). Nous déterminons également les solutions d'entrée minimales pour obtenir la normalisation, lorsque cela est possible. Dans la section 5, nous calculons explicitement les mesures de possibilité et de nécessité correspondantes à la distribution des possibilités de sortie. Tous ces travaux nous permettent de traiter une cascade (section 6) et nous la décrivons par un réseau de neurones min-max. Enfin, nous terminons avec quelques perspectives (section 7).

2 Rappels

Dans cette section, nous rappelons les définitions d'un système à base de règles possibilistes et d'une cascade, ainsi que le système d'équations introduit dans [5]. Nous fixons également quelques notations pour la suite.

2.1 Système à base de règles possibilistes

Un tel système est composé de n règles possibilistes $R^1, R^2, \ldots, R^n$. Pour chaque règle R^i, une matrice de propagation de l'incertitude $\begin{bmatrix} \pi(q_i|p_i) & \pi(q_i|\neg p_i) \\ \pi(\neg q_i|p_i) & \pi(\neg q_i|\neg p_i) \end{bmatrix} = \begin{bmatrix} 1 & s_i \\ r_i & 1 \end{bmatrix}$ encode l'incertitude de "si p_i alors q_i" et de "si $\neg p_i$ alors $\neg q_i$".

Une prémisse non composée p_i est une proposition de la forme "$a_i(x) \in P_i$". L'attribut a_i est appliqué à un élément x. Son information est représentée par une distribution des possibilités $\pi_{a_i(x)} : D_{a_i} \to [0,1]$ définie sur son domaine D_{a_i} et qui est supposée normalisée i.e., $\exists u \in D_{a_i}$ tel que $\pi_{a_i(x)}(u) = 1$. Le degré de possibilité de p_i et celui de sa négation sont définis à l'aide de la mesure de possibilité Π par $\pi(p_i) = \Pi(P_i) = \sup_{u \in P_i} \pi_{a_i(x)}(u)$ et $\pi(\neg p_i) = \Pi(\overline{P_i}) = \sup_{u \in \overline{P_i}} \pi_{a_i(x)}(u)$ respectivement, où P_i est un sous-ensemble de D_{a_i} et $\overline{P_i}$ son complémentaire. On les note respectivement λ_i et ρ_i. Dans le cas d'une prémisse composée $p_i = p_{1,i} \wedge \cdots \wedge p_{k,i}$, on prend $\lambda_i = \min_{j=1}^{k} \pi(p_{j,i})$ et $\rho_i = \max_{j=1}^{k} \pi(\neg p_{j,i})$ [4].

La conclusion q_i est de la forme "$b(x) \in Q_i$", où Q_i est un sous-ensemble de D_b, le domaine de l'attribut b. Le degré de possibilité de q_i et celui de sa négation sont respectivement notés α_i et β_i et sont donnés par $\begin{bmatrix} \pi(q_i) \\ \pi(\neg q_i) \end{bmatrix} = \begin{bmatrix} 1 & s_i \\ r_i & 1 \end{bmatrix} \square_{\min}^{\max} \begin{bmatrix} \lambda_i \\ \rho_i \end{bmatrix}$. Le produit matriciel $\square_{\min}^{\max}$ utilise $\min$ pour le produit et $\max$ pour l'addition. La normalisation de p_i i.e., $\max(\lambda_i, \rho_i) = 1$ entraîne $\alpha_i = \max(s_i, \lambda_i)$ et $\beta_i = \max(r_i, \rho_i)$. Enfin, la distribution des possibilités de l'attribut b associé à la règle R^i est : $\pi^{*i}_{b(x)}(u) = \alpha_i \mu_{Q_i}(u) + \beta_i \mu_{\overline{Q_i}}(u)$ pour tout $u \in D_b$. Dans le cas de n règles, la distribution des possibilités de b est obtenue par la combinaison conjonctive suivante : $\pi^{*}_{b(x)}(u) = \min(\pi^{*1}_{b(x)}(u), \pi^{*2}_{b(x)}(u), \ldots, \pi^{*n}_{b(x)}(u))$.

2.2 Cascade

Une cascade est un système à base de règles possibilistes composée de deux ensembles de règles $R^1, R^2, \ldots, R^n$ et $R'^1, R'^2, \ldots, R'^m$, où les conclusions des R^i et les prémisses des R'^j utilisent le *même attribut*, établissant un chaînage des deux ensembles de règles. La prémisse p'_j de chaque règle R'^j est une proposition "$b(x) \in Q'_j$" où $Q'_j \subseteq D_b$. La conclusion q'_j de R'_j est de la forme "$c(x) \in Q''_j$" où Q''_j est un sous-ensemble de D_c, le domaine de l'attribut c. Les degrés de possibilité associés à R'^j sont calculés de la même manière que ceux des règles R^i : $\lambda'_j = \pi(p'_j)$ et $\rho'_j = \pi(\neg p'_j)$. De même, R'^j est munie d'une matrice de propagation de l'incertitude avec ses paramètres s'_j, r'_j.

2.3 Système d'équations

Dans [5], un système d'équations noté OV $=$ MR $\blacksquare$ IV a été formulé, où OV et IV sont respectivement les vecteurs de sortie et d'entrée. Pour un exemple d'un système de trois règles parallèles, les auteurs ont montré que le système d'équations décrit la distribution des possibilités de sortie et ont proposé de l'utiliser pour effectuer une analyse de sensibilité. Les auteurs de [4] ont explicité le système d'équations pour le cas de deux règles R^1 et R^2 :

$$\begin{bmatrix} \Pi(Q_1 \cap Q_2) \\ \Pi(Q_1 \cap \overline{Q_2}) \\ \Pi(\overline{Q_1} \cap Q_2) \\ \Pi(\overline{Q_1} \cap \overline{Q_2}) \end{bmatrix} = \begin{bmatrix} s_1 & 1 & s_2 & 1 \\ s_1 & 1 & 1 & r_2 \\ 1 & r_1 & s_2 & 1 \\ 1 & r_1 & 1 & r_2 \end{bmatrix} \square^{\min}_{\max} \begin{bmatrix} \lambda_1 \\ \rho_1 \\ \lambda_2 \\ \rho_2 \end{bmatrix}. \quad (1)$$

Le produit matriciel $\square^{\min}_{\max}$ utilise $\max$ pour le produit et $\min$ pour l'addition. Les auteurs ont obtenu $\Pi(Q_1 \cap Q_2) = \min(\alpha_1, \alpha_2)$, $\Pi(Q_1 \cap \overline{Q_2}) = \min(\alpha_1, \beta_2)$, $\Pi(\overline{Q_1} \cap Q_2) = \min(\beta_1, \alpha_2)$ et $\Pi(\overline{Q_1} \cap \overline{Q_2}) = \min(\beta_1, \beta_2)$. Les quatre ensembles utilisés forment une partition de D_b construite à partir des ensembles Q_1, Q_2 utilisés dans les conclusions q_1, q_2 de R^1 et R^2, et de leurs complémentaires.

2.4 Notations

Dans notre travail, toutes les matrices ont leurs coefficients dans $[0, 1]$. A toute matrice $A = [a_{ij}]$, on associe la matrice $A^\circ = [1 - a_{ij}]$. Si A et B sont des matrices de tailles respectives (n, m) et (m, p), la transformation $A \mapsto A^\circ$ permute les deux produits matriciels au sens suivant :

$$(A \square^{\min}_{\max} B)^\circ = A^\circ \square^{\max}_{\min} B^\circ \text{ et } (A \square^{\max}_{\min} B)^\circ = A^\circ \square^{\min}_{\max} B^\circ.$$

Ce résultat ainsi que la propriété $(A^\circ)^\circ = A$ jouent un rôle important pour traiter le cas d'une cascade.

Enfin, nous introduisons un opérateur noté $\boxdot_{\min}$ qui associe à une matrice, la matrice colonne obtenue en prenant le minimum des coefficients de chaque ligne de cette matrice.

3 Système d'équations généralisé

Dans cette section, nous étudions un système composé de n règles possibilistes $R^1, R^2, \ldots, R^n$. Nous introduisons le système d'équations généralisé, que nous notons :

$$O_n = M_n \square^{\min}_{\max} I_n.$$

Pour $n = 2$, notre construction est équivalente à la construction précédemment rappelée, voir (1). Pour comprendre le vecteur de sortie O_n du système d'équations, nous introduisons une partition explicite de D_b. De plus, cette partition est directement liée à une matrice B_n que nous construisons inductivement par rapport au nombre de règles.

3.1 Partition et paramètres

À partir des ensembles $Q_1, Q_2, \ldots, Q_n$ utilisés dans les conclusions des règles et de leurs complémentaires, pour chaque $i = 1, 2, \ldots, n$, on définit $(E_k^{(i)})_{1 \le k \le 2^i}$ une partition de D_b par les deux conditions suivantes :

$$\bullet\ E_1^{(1)} = Q_1 \text{ et } E_2^{(1)} = \overline{Q_1} \quad (2a)$$

et pour $i > 1$:

$$\bullet\ E_k^{(i)} = \begin{cases} E_k^{(i-1)} \cap Q_i & \text{if } 1 \le k \le 2^{i-1} \\ E_{k-2^{i-1}}^{(i-1)} \cap \overline{Q_i} & \text{if } 2^{i-1} < k \le 2^i \end{cases}. \quad (2b)$$

Pour $i = 1, 2, \ldots, n$, on définit les matrices M_i, I_i et B_i selon :

— $s_1, s_2, \ldots, s_i$ et $r_1, r_2, \ldots, r_i$ pour M_i,
— $\lambda_1, \lambda_2, \ldots, \lambda_i$ et $\rho_1, \rho_2, \ldots, \rho_i$ pour I_i,
— $\alpha_1, \alpha_2, \ldots, \alpha_i$ et $\beta_1, \beta_2, \ldots, \beta_i$ pour B_i.

3.2 Construction de M_i

Pour $i = 1$, on prend $M_1 = \begin{bmatrix} s_1 & 1 \\ 1 & r_1 \end{bmatrix}$. Pour $i > 1$, on définit M_i de taille $(2^i, 2i)$ par une construction de matrices par blocs :

$$M_i = \left[\begin{array}{c|c} M_{i-1} & S_i \\ \hline M_{i-1} & R_i \end{array}\right].$$

où $S_i = \begin{bmatrix} s_i & 1 \\ s_i & 1 \\ \vdots & \vdots \\ s_i & 1 \end{bmatrix}$ et $R_i = \begin{bmatrix} 1 & r_i \\ 1 & r_i \\ \vdots & \vdots \\ 1 & r_i \end{bmatrix}$ sont de taille $(2^{i-1}, 2)$.

On note $N_1, N_2, \ldots, N_{2^i}$ les lignes de M_i.

3.3 Construction de I_i

Pour $i = 1$, on prend $I_1 = \begin{bmatrix} \lambda_1 \\ \rho_1 \end{bmatrix}$. Pour $i > 1$, on définit I_i de taille $(2i, 1)$:

$$I_i = \begin{bmatrix} I_{i-1} \\ \hline \lambda_i \\ \rho_i \end{bmatrix} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_{2i} \end{bmatrix}.$$

où $\theta_{2j-1} = \lambda_j$ et $\theta_{2j} = \rho_j$ pour $j = 1, 2, \ldots, i$.

3.4 Construction de B_i

Pour $i = 1$, on prend $B_1 = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix}$. Pour $i > 1$, on définit B_i de taille $(2^i, i)$ par une construction de matrices par blocs :

$$B_i = \left[\begin{array}{c|c} & \alpha_i \\ & \alpha_i \\ B_{i-1} & \vdots \\ & \alpha_i \\ \hline & \beta_i \\ & \beta_i \\ B_{i-1} & \vdots \\ & \beta_i \end{array}\right].$$

Les lignes $L_1, L_2, \ldots, L_{2^i}$ de B_i s'obtiennent à partir des lignes $L'_1, L'_2, \ldots, L'_{2^{i-1}}$ de B_{i-1} par la règle suivante :

$$L_k = \begin{cases} (L'_k, \alpha_i) & \text{if} \quad 1 \leq k \leq 2^{i-1} \\ (L'_{k-2^{i-1}}, \beta_i) & \text{if} \quad 2^{i-1} < k \leq 2^i \end{cases}. \quad (3)$$

Nous pouvons retrouver la matrice M_i à partir de la matrice B_i de manière explicite.

3.5 Relation entre B_i et la partition

Soit $k \in \{1, 2, \ldots, 2^i\}$ et $L_k = (\gamma_1, \gamma_2, \ldots, \gamma_i)$ une ligne de la matrice B_i avec $\gamma \in \{\alpha, \beta\}$. Alors l'ensemble correspondant $E_k^{(i)}$ de la partition est égal à :

$$E_k^{(i)} = T_1 \cap T_2 \cdots \cap T_i \text{ où } T_j = \begin{cases} Q_j \text{ si } \gamma_j = \alpha_j \\ \overline{Q_j} \text{ si } \gamma_j = \beta_j \end{cases}. \quad (4)$$

Ce résultat étant clair pour $i = 1$, il se déduit de la description des lignes de B_i par les lignes de B_{i-1}, voir (3).

3.6 Coefficients de $\boxdot_{\min} B_i$

Pour tout $i = 1, 2, \ldots, n$, on pose :

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

Pour tout $k \in \{1, 2, \ldots, 2^i\}$, nous avons les relations suivantes entre les coefficients de $\boxdot_{\min} B_i$ et ceux de $\boxdot_{\min} B_{i-1}$:

$$o_k^{(i)} = \begin{cases} \min(o_k^{(i-1)}, \alpha_i) & \text{si } 1 \leq k \leq 2^{i-1} \\ \min(o_{k-2^{i-1}}^{(i-1)}, \beta_i) & \text{si } 2^{i-1} < k \leq 2^i \end{cases}. \quad (5)$$

Ce résultat est directement déduit de (3) et de l'associativité de la fonction $\min$. On a :

Théorème 1 *Le produit matriciel min-max de M_i par la matrice colonne I_i est obtenu en appliquant l'opérateur $\boxdot_{min}$ à la matrice B_i :*

$$M_i \Box_{\max}^{\min} I_i = \boxdot_{\min} B_i. \quad (6)$$

La preuve abrégée de ce résultat est la suivante. Elle est claire pour $i = 1$. En partant de l'hypothèse (6) pour $i - 1$, on montre que chaque coefficient $o_k^{(i)}$ de $\boxdot_{\min} B_i$ est égal à $N_k \Box_{\max}^{\min} I_i$, N_k étant la k-ième ligne de M_i.

4 Propriétés du système d'équations

Dans cette section, nous étudions les propriétés du système d'équations. Nous établissons une formule additive pour $\pi^*_{b(x)}$ à partir de la partition $(E_k^{(i)})_{1 \leq k \leq 2^i}$. Nous donnons une condition nécessaire et suffisante pour la normalisation de $\pi^*_{b(x)}$. Nous montrons ensuite que, en

supprimant les ensembles vides de la partition et les lignes correspondantes de O_i, M_i et B_i, nous obtenons des matrices $\mathcal{O}_i$, $\mathcal{M}_i$ et $\mathcal{B}_i$ avec un nombre raisonnable de lignes. Nous étudions également les solutions extrémales pour obtenir la normalisation et donnons une méthode pour déduire le système d'équations obtenu par suppression d'une règle.

4.1 Formule additive pour $\pi^*_{b(x)}$

En se basant sur l'égalité $O_i = \boxdot_{\min} B_i$, voir (6), et en notant $\mu_{E_1^{(i)}}, \mu_{E_2^{(i)}}, \ldots, \mu_{E_{2^i}^{(i)}}$ les fonctions caractéristiques des ensembles $E_1^{(i)}, E_2^{(i)}, \ldots, E_{2^i}^{(i)}$, on obtient :

Théorème 2 *La distribution des possibilités de sortie $\pi^*_{b(x),i}$ associée aux i premières règles est :*

$$\pi^*_{b(x),i} = \sum_{1 \leq k \leq 2^i} o_k^{(i)} \mu_{E_k^{(i)}}. \qquad (7)$$

Cette égalité *fonctionnelle*, où le membre de droite est une combinaison linéaire de *fonctions caractéristiques* d'ensembles d'une partition, se démontre par récurrence sur i en utilisant (2), (5) et l'associativité de la fonction $\min$. Il s'ensuit que $\forall u \in D_b$, il existe un unique indice k_0 tel que $u \in E_{k_0}^{(i)}$ et $\pi^*_{b(x),i}(u) = o_{k_0}^{(i)}$. Il en résulte que $\pi^*_{b(x),i}$ est normalisée si et seulement si :

$$\exists k \in \{1, 2, \ldots, 2^i\} \text{ tq } E_k^{(i)} \neq \emptyset \text{ et } o_k^{(i)} = 1. \qquad (8)$$

On déduit de (7) que la mesure de possibilité de chaque ensemble *non vide* $E_k^{(i)}$ de la partition est égale à $o_k^{(i)}$. Il est alors naturel d'introduire :

$$J = \left\{ k \in \{1, 2, \ldots, 2^i\} \mid E_k^{(i)} \neq \emptyset \right\}.$$

Posons $\omega = \text{card}(J)$. On a $\omega \leq \min(d, 2^i)$, où $d = \text{card}(D_b)$. En rangeant les entiers de J en une suite strictement croissante : $1 \leq k_1 < k_2 < \cdots < k_\omega \leq 2^i$, on déduit :

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

Dans ce qui suit, on note $\mathcal{O}_i$, $\mathcal{M}_i$ et $\mathcal{B}_i$, les matrices obtenues à partir de O_i, M_i et B_i respectivement, en supprimant chaque ligne dont l'indice n'est pas dans J.

4.2 Solutions de $\pi^*_{b(x),i}(u) = 1$

Dans ce qui suit, nous étudions comment obtenir $\pi^*_{b(x),i}(u) = 1$ pour une valeur $u \in D_b$. En combinant (6) et (7), on a $\pi^*_{b(x),i}(u) = N \square^{\min}_{\max} I_i$, où $N = [t_1, t_2, \ldots, t_{2i-1}, t_{2i}]$ est une ligne de $\mathcal{M}_i$ et $t_j \in \{1, r_*, s_*\}$.

Avec les notations de la sous-section 3.3, $\pi^*_{b(x),i}(u) = \min_{1 \leq j \leq 2i} \max(t_j, \theta_j)$. Donc $\pi^*_{b(x),i}(u) = 1$ est équivalent à $\forall j \in E$, $\theta_j = 1$, où $E = \{j \in \{1, 2, \ldots, 2i\} \mid t_j < 1\}$. Ainsi, la normalisation de la distribution des possibilités de b peut être établie par la résolution d'un système min-max de ω équations, où le second membre d'au moins une équation doit être égal à 1. Par conséquent, il est intéressant d'étudier s'il existe des solutions extrémales I_i pour obtenir $\pi^*_{b(x),i}(u) = 1$.

Dans ce qui suit, nous notons $X = [x_j]_{1 \leq j \leq 2i}$ un vecteur d'entrée vérifiant l'hypothèse de normalisation : $\forall k \in \{1, 2, \ldots, i\}$: $\max(x_{2k-1}, x_{2k}) = 1$, satisfaisant l'équation suivante :

$$N \square^{\min}_{\max} X = 1. \qquad (9)$$

Pour deux vecteurs $X = [x_j]_{1 \leq j \leq 2i}$ et $X' = [x'_j]_{1 \leq j \leq 2i}$, on pose :

$$X \leq X', \text{ ssi } \forall j \in \{1, 2, \ldots, 2i\} : x_j \leq x'_j.$$

Nous pouvons étudier l'existence d'une solution minimale unique notée S_{min} ou d'une solution maximale $S_{\max}$.

Il est clair que $S_{max} = [\theta_j^*]$ où $\theta_j^* = 1$ pour chaque $j = 1, 2, \ldots, 2i$. Pour cette solution maximale S_{max}, on remarque que $\forall k \in \{1, 2, \ldots, i\}, (\lambda_k, \rho_k) = (1, 1)$: pour chaque règle, la prémisse est considérée comme inconnue [4].

Cherchons une unique solution minimale :
• Si on a :

$$\exists k \in \{1, 2, \ldots, i\} \text{ tq } (t_{2k-1}, t_{2k}) = (1, 1), \qquad (10)$$

alors l'équation (9) n'admet pas une solution minimale.

• Si on suppose que :

$$\forall k \in \{1, 2, \ldots, i\} \text{ on a } (t_{2k-1}, t_{2k}) \neq (1, 1), \quad (11)$$

alors il existe une solution minimale unique $S_{\min} = [\theta^*_j]_{1 \leq j \leq 2i}$ où $(\theta^*_{2k-1}, \theta^*_{2k}) = \begin{cases} (1,0) & \text{if } t_{2k} = 1 \\ (0,1) & \text{if } t_{2k-1} = 1 \end{cases}$. Ainsi, avec l'hypothèse (11), nous donnons une condition suffisante pour la normalisation de $\pi^*_{b(x),i}$.

4.3 Reconstruire le système d'équations obtenu par suppression d'une règle

Pour effectuer de l'analyse de sensibilité de notre système de i règles $R^1, R^2, \ldots, R^i$, on se propose de supprimer une règle R^z et d'étudier le système d'équations correspondant. Notons $B_i^{\not z}$ la matrice associée à ce système de $i-1$ règles par la construction de la sous-section 3.4. Nous affirmons que la matrice $B_i^{\not z}$, qui est de taille $(2^{i-1}, i-1)$, peut être obtenue à partir de la matrice B_i par la règle pratique suivante :

1. Nous supprimons la z-ième colonne de B_i. On obtient alors une matrice de taille $(2^i, i-1)$, où chaque ligne se répète une et une seule fois.
2. Dans la matrice résultante, nous supprimons les lignes $L_{k'}$ dans toutes les paires de lignes $(L_k, L_{k'})$ où $L_k = L_{k'}$ et $k < k'$.

Après ces deux opérations, on obtient $B_i^{\not z}$. On peut alors déduire la partition $(E_k^{\not z})_{1 \leq k \leq 2^{i-1}}$ de D_b à partir de $B_i^{\not z}$ en utilisant la relation (4). Ses ensembles peuvent également être déterminés directement à partir de la partition $(E_k^{(i)})_{1 \leq k \leq 2^i}$. En effet, pour tout $k \in \{1, 2, \ldots, 2^{i-1}\}$, on peut trouver deux indices $k', k'' \in \{1, 2, \ldots, 2^i\}$ tels que :

$$E_k^{\not z} = E_{k'}^{(i)} \cup E_{k''}^{(i)} \quad (12)$$

où, avec les notations de (4), $E_{k'}^{(i)}$ et $E_{k''}^{(i)}$ ne diffèrent que sur la composante T_z, par exemple $T_z = Q_z$ pour $E_{k'}^{(i)}$, et $T_z = \overline{Q}_z$ pour $E_{k''}^{(i)}$. Si $E_k^{\not z} \neq \emptyset$, la décomposition (12) est unique, où $E_{k'}^{(i)}$ et $E_{k''}^{(i)}$ satisfont toujours l'hypothèse ci-dessus. Ces deux ensembles peuvent être facilement déterminés de manière algorithmique.

Enfin, nous déduisons de la relation (12) que $\omega^{\not z} = \text{card}(J^{\not z}) \leq \omega$ avec $J^{\not z} = \left\{k \in \{1, 2, \ldots, 2^{i-1}\} \;\middle|\; E_k^{\not z} \neq \emptyset\right\}$. La partition $(E_k^{\not z})_{k \in J^{\not z}}$ et la matrice $\mathcal{B}_i^{\not z}$ sont particulièrement utiles pour effectuer de l'analyse de sensibilité.

5 Mesures de possibilité et de nécessité associées à $\pi^*_{b(x)}$

Rappelons que la mesure de possibilité Π^* et la mesure de nécessité N^* associées à $\pi^*_{b(x)}$ sont définies par [2] :

$$\Pi^* : \mathcal{P}(D_b) \to [0,1] : Q \mapsto \Pi^*(Q) = \max_{u \in Q} \pi^*_{b(x)}(u),$$

$$N^* : \mathcal{P}(D_b) \to [0,1] : Q \mapsto N^*(Q) = 1 - \Pi^*(\overline{Q}).$$

Pour une proposition p : "$b(x) \in Q$", le degré de possibilité $\pi(p)$ et son degré de nécessité $n(p)$ sont donnés par [2] :

$$\pi(p) = \Pi^*(Q) \text{ et } n(p) = 1 - \pi(\neg p) = N^*(Q).$$

Dans ce qui suit, nous déduisons de (7) des formules explicites pour $\Pi^*(Q)$ et $N^*(Q)$ où nous utilisons une fonction ε qui vérifie si un ensemble n'est pas vide :

$$\varepsilon(T) = \begin{cases} 1 & \text{si } T \neq \emptyset \\ 0 & \text{if } T = \emptyset \end{cases}.$$

5.1 Mesure de possibilité

Pour tout sous-ensemble $Q \subseteq D_b$, on a $Q = \bigcup_{1 \leq i \leq \omega} E_{k_i}^{(n)} \cap Q$, et nous savons que $\forall u \in E_{k_i}^{(n)}$ nous avons $\pi^*_{b(x)}(u) = o_{k_i}^{(n)}$. On obtient donc la mesure de possibilité de Q par :

$$\Pi^*(Q) = \max_{u \in Q} \pi^*_{b(x)}(u) = \max_{1 \leq i \leq \omega \text{ s.t } E_{k_i}^{(n)} \cap Q \neq \emptyset} o_{k_i}^{(n)}.$$

Par conséquent, nous pouvons reformuler ce résultat comme suit :

$$\begin{aligned} \Pi^*(Q) &= \max_{1 \leq i \leq \omega} \varepsilon(E_{k_i}^{(n)} \cap Q) \dot{o}_{k_i}^{(n)} \\ &= \max_{1 \leq i \leq \omega} \min(\varepsilon(E_{k_i}^{(n)}) \cap Q), o_{k_i}^{n)}). \quad (13) \end{aligned}$$

Soit Δ_Q la matrice de taille $(1, \omega)$ définie par :

$$\Delta_Q = \left[\varepsilon(E_{k_1}^{(n)} \cap Q) \quad \varepsilon(E_{k_2}^{(n)} \cap Q) \quad \dots \quad \varepsilon(E_{k_\omega}^{(n)} \cap Q)\right].$$

Alors, l'égalité (13) est exactement :

$$\Pi^*(Q) = \Delta_Q \square_{\min}^{\max} \mathcal{O}_n. \tag{14}$$

5.2 Mesure de nécessité

En utilisant (14), nous avons pour $\overline{Q}$:

$$\Pi^*(\overline{Q}) = \Delta_{\overline{Q}} \square_{\min}^{\max} \mathcal{O}_n.$$

La mesure de nécessité est alors :

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

Par la relation entre $\square_{\max}^{\min}$ et $\square_{\min}^{\max}$ on obtient :

$$N^*(Q) = (\Delta_{\overline{Q}} \square_{\min}^{\max} \mathcal{O}_n)^\circ = \Delta_{\overline{Q}}{}^\circ \square_{\max}^{\min} \mathcal{O}_n{}^\circ. \tag{15}$$

6 Cascade

Dans cette section, nous utilisons deux ensembles de règles possibilistes : les n règles $R^1, R^2, \dots, R^n$ et les m règles $R'^1, R'^2, \dots, R'^m$. Nous formons $\mathcal{O}_n = \mathcal{M}_n \square_{\max}^{\min} I_n$ pour le premier ensemble de règles et $\mathcal{O}'_m = \mathcal{M}'_m \square_{\max}^{\min} I'_m$ pour le second, où nous notons respectivement ω et ω' le nombre d'ensembles non vides des partitions $(E_k^{(n)})$ et $(E_k^{'(m)})$. Dans ce qui suit, nous établissons une relation liant les deux systèmes d'équations et associons à une telle cascade un réseau de neurones min-max.

6.1 vecteurs I'_m et O'_m

Chaque prémisse p'_j d'une règle R'_j est une proposition de la forme "$b(x) \in Q'_j$". Par conséquent, nous obtenons λ'_j et ρ'_j par le calcul des mesures de possibilité de Q'_j et $\overline{Q'_j}$:

$$\lambda'_j = \Pi^*(Q'_j) \text{ et } \rho'_j = \Pi^*(\overline{Q'_j}).$$

Par l'égalité (14), on définit le vecteur d'entrée I'_m de taille $(2m, 1)$, comme un produit max-min entre une matrice $\Delta = [\delta_{t,k}]_{1 \le t \le 2m, 1 \le k \le \omega}$, et $\mathcal{O}_n$ de taille $(\omega, 1)$:

$$I'_m = \Delta \square_{min}^{max} \mathcal{O}_n. \tag{16}$$

Les coefficients de Δ sont 1 ou 0, et s'obtiennent comme suit :

$$\delta_{t,k} = \begin{cases} \varepsilon(E_k^{(n)} \cap Q'_j) & \text{if } t = 2j-1 \text{ et } 1 \le j \le m \\ \varepsilon(E_k^{(n)} \cap \overline{Q'_j}) & \text{if } t = 2j \text{ and } 1 \le j \le m \end{cases}.$$

Nous notons $S_1, S_2, \dots, S_{2m}$ les lignes de Δ. Chaque ligne de Δ est en fait la matrice $\Delta_{Q'}$ associée à un ensemble $Q' \subseteq D_b$ comme dans (14) :

$$\Delta = \begin{bmatrix} \Delta_{Q'_1} \\ \Delta_{\overline{Q'_1}} \\ \vdots \\ \Delta_{Q'_m} \\ \Delta_{\overline{Q'_m}} \end{bmatrix}.$$

Le vecteur de sortie du second système $\mathcal{O}'_m$ s'obtient à partir du premier système, Δ et $\mathcal{M}'_m$:

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

6.2 Représentation par un réseau de neurones min-max

Nous établissons une correspondance entre une cascade et un réseau de neurones min-max, qui donne lieu à une représentation graphique de la construction d'une cascade. Dubois et Prade ont suggéré que le système d'équations associé à une cascade aurait une ressemblance structurelle avec un tel réseau [4].

Les matrices $I_n{}^\circ$, $\mathcal{M}_n{}^\circ$ et $\mathcal{M}'_m{}^\circ$ permettent d'exprimer les équations impliquées dans la cascade en utilisant uniquement l'opérateur $(A \square_{\min}^{\max} B)^\circ$: $\mathcal{O}_n = (\mathcal{M}_n{}^\circ \square_{\min}^{\max} I_n{}^\circ)^\circ$, $I'_m{}^\circ = (\Delta \square_{\min}^{\max} \mathcal{O}_n)^\circ$ et $\mathcal{O}'_m = (\mathcal{M}'_m{}^\circ \square_{\min}^{\max} I'_m{}^\circ)^\circ$.

Nous définissons le réseau de neurones min-max à quatre couches comme suit :

- la couche 1 possède $2n$ neurones d'entrée $i_1, i_2, \dots, i_{2n}$,
- la couche 2 possède ω neurones cachés $h_1, h_2, \dots, h_\omega$,

— la couche 3 possède $2m$ neurones cachés $h'_1, h'_2, \ldots, h'_{2m}$,
— la couche 4 a ω' neurones de sortie : $o'_1, o'_2, \ldots, o'_{\omega'}$.

Dans ce réseau de neurones, la valeur d'entrée de chaque neurone s'obtient par l'opérateur $(A \square^{\max}_{\min} B)^\circ$. Sa valeur de sortie est donnée par la fonction d'activation $f(x) = x$.

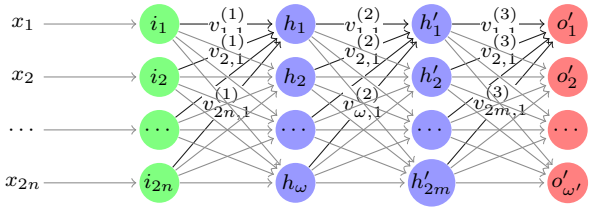


Figure 1 – Architecture du réseau de neurones

Nous explicitons son architecture (Figure 1) et définissons les arêtes :

— $x_j = 1 - \theta_j$, un coefficient de ${I_n}^\circ$ avec $1 \leq j \leq 2n$,
— $v^{(1)}_{i,j} = 1 - t_{j,i}$, un coefficient de ${\mathcal{M}_n}^\circ$ avec $1 \leq j \leq \omega$ et $1 \leq i \leq 2n$,
— $v^{(2)}_{i,j} = \delta_{j,i}$, un coefficient de Δ avec $1 \leq j \leq 2m$ et $1 \leq i \leq \omega$,
— $v^{(3)}_{i,j} = 1 - t'_{j,i}$, un coefficient de ${\mathcal{M}'_m}^\circ$ avec $1 \leq j \leq \omega'$ et $1 \leq i \leq 2m$.

Ce réseau de neurones présente une certaine ressemblance avec un réseau de neurones flou hybride [1]. En utilisant davantage de couches, nous pouvons étendre ce réseau de neurones pour prendre en compte les calculs de λ, ρ lorsque les prémisses sont composées.

7 Conclusion et perspectives

Dans cet article, nous avons donné une construction canonique pour les matrices gouvernant le système d'équations min-max de Farreny et Prade [5]. Nous avons obtenu une formule explicite pour la distribution des possibilités de sortie et calculé les mesures de possibilité et de nécessité correspondantes. Enfin, nous avons montré que le système d'équations formé pour une cascade peut être représenté par un réseau de neurones min-max. Le système d'équations de Farreny et Prade peut être utilisé pour de nombreuses études. Nous pouvons effectuer une analyse de sensibilité, comme l'avait initialement suggéré Farreny et Prade. Nous pouvons rechercher des conditions générales pour établir qu'une base de règles possibiliste est cohérente. Ces conditions porteraient sur les paramètres des règles et le vecteur d'entrée. Enfin, pour l'apprentissage possibiliste, il serait intéressant d'adapter une méthode de descente de gradient pour notre réseau de neurones min-max.

Références

[1] James J BUCKLEY et Yoichi HAYASHI : Fuzzy neural networks : A survey. *Fuzzy sets and systems*, 66(1):1–13, 1994.

[2] D. DUBOIS et H. PRADE : *Possibility theory : an approach to computerized processing of uncertainty*. Plenum Press, New York, 1988. ISBN 0306425203.

[3] D. DUBOIS et H. PRADE : Possibility theory and its applications : Where do we stand ? *In Handbook of Computational Intelligence*, 2015.

[4] Didier DUBOIS et Henri PRADE : From possibilistic rule-based systems to machine learning-a discussion paper. *In International Conference on Scalable Uncertainty Management*, pages 35–51. Springer, 2020.

[5] H FARRENY et H PRADE : Explications de raisonnements dans l'incertain. *Revue d'intelligence artificielle*, 4(2):43–75, 1990.

[6] Henri FARRENY, Henri PRADE et E WYSS : Approximate reasoning in a rule-based expert system using possibility theory : A case study. *In IFIP Congress*, pages 407–414, 1986.

[7] Lotfi Asker ZADEH : Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems*, 1(1):3–28, 1978.