



**HAL**  
open science

# Apprentissage des paramètres des règles d'un système à base de règles possibilistes

Ismail Baaj

► **To cite this version:**

Ismail Baaj. Apprentissage des paramètres des règles d'un système à base de règles possibilistes. Rencontres francophones sur la logique floue et ses applications, Oct 2022, Toulouse, France. pp.199-206. <hal-03859322>

**HAL Id: hal-03859322**

**<https://hal.science/hal-03859322v1>**

Submitted on 30 Nov 2023

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.



HAL Authorization

# Apprentissage des paramètres des règles d'un système à base de règles possibilistes

Ismaïl Baaj

LIP6, Sorbonne Université, Paris, France.

ismail.baaj@lip6.fr

## Résumé :

Dans cet article, nous introduisons un paradigme d'apprentissage des paramètres des règles d'un système à base de règles possibilistes en fonction d'une donnée d'entraînement. Pour tout système composé de  $n$  règles possibilistes parallèles, nous introduisons un système d'équations noté  $(\Sigma_n)$ , analogue au système d'équations de Farreny-Prade, où l'inconnue est un vecteur dont les composantes sont les paramètres des règles qu'on doit déterminer en fonction d'une donnée d'entraînement. Nous établissons les conditions nécessaires et suffisantes pour que le système  $(\Sigma_n)$  admette des solutions. Dans ce cas, on montre qu'il existe une unique solution maximale et une unique solution minimale. Enfin, nos résultats sont illustrés par un exemple.

## Mots-clés :

Théorie des possibilités, apprentissage, système à base de règles, systèmes d'équations floues

## Abstract:

In this paper, we introduce a paradigm for learning the parameters of the rules of a possibilistic rule-based system according to a training datum. For any system composed of  $n$  parallel possibilistic rules, we introduce a system of equations denoted  $(\Sigma_n)$ , analogous to the Farreny-Prade equation system, where the unknown is a vector whose components are the parameters of the rules that must be determined according to a training datum. We establish the necessary and sufficient conditions for the system  $(\Sigma_n)$  to have solutions. In this case, we show that there is a unique maximal and a unique minimal solution. Finally, our results are illustrated by an example.

## Keywords:

possibility theory, learning, rule-based system, systems of fuzzy equations

## 1 Introduction

La théorie des possibilités traite les informations incomplètes ou imprécises [3, 4] à l'aide de deux mesures duales appelées possibilité et nécessité. Ces mesures permettent de distinguer ce qui est possible sans être tout à fait certain et ce qui est certain dans une moindre mesure. Dans [5], Dubois et Prade ont récemment mis l'accent sur le développement de méthodes d'apprentissage possibilistes qui seraient compatibles avec le raisonnement basé sur des

règles. Dans ce but, ces auteurs ont rappelé le système d'équations min-max de Farreny et Prade [7], qui a été proposé pour développer les capacités explicatives des systèmes à base de règles possibilistes. Ce système d'équations est composé d'un vecteur d'entrée qui contient les degrés de possibilité des prémisses des règles, d'une matrice contenant les paramètres des règles, qui sont des degrés de possibilité conditionnels entre la prémisse et la conclusion de chaque règle, et d'un vecteur de sortie, qui décrit la distribution des possibilités de sortie d'une inférence. Dubois et Prade ont souligné que le système d'équations associé à une cascade, i.e, un système utilisant deux ensembles chaînés de règles possibilistes parallèles, peut être décrit par un réseau de neurones min-max [5].

Suite à ces travaux, les auteurs de [2] ont récemment donné une construction canonique des matrices régissant le système d'équations de Farreny-Prade (voir aussi [1]). Ils ont étendu ce système d'équations au cas d'une cascade en établissant une relation entrée-sortie entre les deux systèmes d'équations associés à chaque ensemble de règles possibilistes parallèles. Ils ont montré également qu'une cascade peut être représentée par un réseau de neurones explicite de type min-max. Cependant, les auteurs ne fournissent pas de méthode pour l'apprentissage des poids du réseau, qui sont les paramètres associés à chaque règle. Nous montrons, dans cet article, que l'apprentissage de ces paramètres en fonction des données d'entraînement peut se réaliser par la résolution d'un système d'équations déduit du système initial. Dans notre cadre, une donnée d'entraînement

est une paire constituée d'un vecteur d'entrée et d'un vecteur de sortie du système d'équations initial, qui sont instanciés. Les composantes du vecteur d'entrée sont les degrés de possibilité des prémisses des règles et le vecteur de sortie décrit une distribution des possibilités de sortie ciblée. Les degrés de possibilité associés à chaque prémisses de règle sont supposés être normalisés, i.e, pour une règle donnée, le degré de possibilité de sa prémisses ou celui de sa négation, est égal à un.

L'article est structuré comme suit. Nous commençons par rappeler le système d'équations de Farreny et Prade [7] associé à un système de règles possibilistes et sa généralisation par [2] (Section 2). Dans la section 3, à un ensemble de  $n$  règles possibilistes parallèles, nous associons un système d'équations noté  $(\Sigma_n)$ , qui est analogue à celui de Farreny et Prade. Étant donné une donnée d'entraînement, nous montrons que les solutions de ce nouveau système d'équations sont des vecteurs dont les composantes attribuent une valeur cohérente à chaque paramètre d'une règle. Dans la section 4, nous énonçons le résultat principal, qui est formulé dans le Théorème 1 : nous déterminons l'ensemble des solutions de  $(\Sigma_n)$ , on dit alors que le système  $(\Sigma_n)$  est compatible. Enfin, nous illustrons nos constructions par un exemple (Section 5) et concluons par quelques perspectives.

## 2 Rappels

Dans cette section, nous rappelons la définition d'un système à base de règles possibilistes parallèles et, en suivant [2], le système d'équations généralisé associé à ce système. Nous montrons une propriété utile du vecteur de sortie de ce système d'équations qui est déduite de la normalisation des degrés de possibilité des prémisses des règles.

### 2.1 Système à base de règles possibilistes

Un système à base de règles possibilistes est composé d'un ensemble de  $n$  règles possibi-

listes parallèles  $R^1, R^2, \dots, R^n$ . Chaque règle  $R^i$  est de la forme : "si  $p_i$  alors  $q_i$ ", où  $p_i$  est sa prémisses et  $q_i$  sa conclusion. La règle est munie d'une matrice de propagation de l'incertitude  $\begin{bmatrix} \pi(q_i|p_i) & \pi(q_i|\neg p_i) \\ \pi(\neg q_i|p_i) & \pi(\neg q_i|\neg p_i) \end{bmatrix} = \begin{bmatrix} 1 & s_i \\ r_i & 1 \end{bmatrix}$ ;  $s_i$  et  $r_i$  sont appelés les paramètres de la règle, voir [5, 6].

La prémisses  $p_i$  est une proposition de la forme " $a_i(x) \in P_i$ ". L'attribut  $a_i$  est appliqué à un élément  $x$  et son information est représentée par une distribution des possibilités  $\pi_{a_i(x)} : D_{a_i} \rightarrow [0, 1]$ , où  $D_{a_i}$  est le domaine de l'attribut  $a_i$ . Nous supposons que la distribution des possibilités  $\pi_{a_i(x)}$  est normalisée c'est-à-dire, il existe  $u \in D_{a_i}$  tel que  $\pi_{a_i(x)}(u) = 1$ . Les degrés de possibilité de la prémisses  $p_i$  et de celui de sa négation sont définis par la mesure de possibilité  $\Pi$  déduite de  $\pi_{a_i(x)}$  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 complément. On les note respectivement  $\lambda_i$  et  $\rho_i$  et on a  $\max(\lambda_i, \rho_i) = 1$ .

Si la prémisses est composée, i.e,  $p_i = p_{1,i} \wedge \dots \wedge p_{k,i}$ , on pose  $\lambda_i = \min_{j=1}^k \pi(p_{j,i})$  et  $\rho_i = \max_{j=1}^k \pi(\neg p_{j,i})$ , voir [5], et nous avons toujours  $\max(\lambda_i, \rho_i) = 1$ .

La conclusion  $q_i$  est une proposition de la forme " $b(x) \in Q_i$ ", où  $b$  est l'attribut de sortie et  $Q_i$  est un sous-ensemble de son domaine  $D_b$ . Le degré de possibilité de  $q_i$  et celui de sa négation sont respectivement notés  $\alpha_i$  et  $\beta_i$ . Ils sont donnés par :

$$\begin{bmatrix} \pi(q_i) \\ \pi(\neg q_i) \end{bmatrix} = \begin{bmatrix} 1 & s_i \\ r_i & 1 \end{bmatrix} \square_{\min}^{\max} \begin{bmatrix} \lambda_i \\ \rho_i \end{bmatrix},$$

où l'opérateur  $\square_{\min}^{\max}$  utilise  $\min$  comme produit et  $\max$  comme addition. En conséquence de  $\max(\lambda_i, \rho_i) = 1$ , nous avons :

$$\alpha_i = \max(s_i, \lambda_i), \quad (1a)$$

$$\beta_i = \max(r_i, \rho_i), \quad (1b)$$

$$\text{et } \max(\alpha_i, \beta_i) = 1. \quad (1c)$$

Enfin, la distribution des possibilités de l'attribut  $b$  associé à la règle  $R^i$  est :

$$\pi_{b(x)}^{*i}(u) = \alpha_i \mu_{Q_i}(u) + \beta_i \mu_{\overline{Q}_i}(u) \forall u \in D_b,$$

où  $\mu_{Q_i}, \mu_{\overline{Q}_i}$  sont les fonctions caractéristiques de  $Q_i$  et  $\overline{Q}_i$ , respectivement. Avec  $n$  règles, la distribution des possibilités de sortie est obtenue par une combinaison conjonctive basée sur min :

$$\pi_{b(x)}^*(u) = \min(\pi_{b(x)}^{*1}(u), \pi_{b(x)}^{*2}(u), \dots, \pi_{b(x)}^{*n}(u)).$$

## 2.2 Système d'équations

Dans [7], un système d'équations noté OV = MR ■ IV a été formulé, où OV et IV sont respectivement nommés les vecteurs de sortie et d'entrée. Dans le cas de  $n$  règles, une construction canonique des matrices régissant ce système d'équations a été donnée par [2]. Le système d'équations résultant est noté  $O_n = M_n \square_{\max}^{\min} I_n$ . L'opérateur  $\square_{\max}^{\min}$  utilise max comme produit et min comme addition. Les matrices  $O_n, M_n$  et  $I_n$  sont de taille  $(2^n, 1)$ ,  $(2^n, 2n)$  et  $(2n, 1)$  respectivement.

Pour  $i = 1, 2, \dots, n$ , les composantes du vecteur d'entrée  $I_i = [\theta_j]_{1 \leq j \leq 2i}$  sont construites selon les degrés de prémisses  $\lambda_1, \lambda_2, \dots, \lambda_i$  et  $\rho_1, \rho_2, \dots, \rho_i$ . La matrice  $M_i = [m_{kj}]_{1 \leq k \leq 2^i, 1 \leq j \leq 2i}$  est associée aux paramètres des règles  $s_1, s_2, \dots, s_i$  et  $r_1, r_2, \dots, r_i$ , et elle est également notée  $M_i(s_i, r_i)$ .

• Pour  $i = 1$ , on prend  $I_1 = \begin{bmatrix} \lambda_1 \\ \rho_1 \end{bmatrix}$  et  $M_1 = \begin{bmatrix} s_1 & 1 \\ 1 & r_1 \end{bmatrix}$ .

• Pour  $i > 1$ ,  $I_i = \begin{bmatrix} I_{i-1} \\ \lambda_i \\ \rho_i \end{bmatrix}$  est de taille  $(2^i, 1)$  et on définit  $M_i$  de taille  $(2^i, 2i)$  par la construction de matrices par blocs suivante :

$$M_i = \left[ \begin{array}{c|c} M_{i-1} & S_i \\ \hline M_{i-1} & R_i \end{array} \right], \text{ où } S_i = \begin{bmatrix} s_i & 1 \\ s_i & 1 \\ \vdots & \vdots \\ s_i & 1 \end{bmatrix} \text{ et } R_i =$$

$$\begin{bmatrix} 1 & r_i \\ 1 & r_i \\ \vdots & \vdots \\ 1 & r_i \end{bmatrix} \text{ sont de tailles } (2^{i-1}, 2).$$

Dans le vecteur de sortie  $O_i = [o_k^{(i)}]_{1 \leq k \leq 2^i}$ , pour  $i = 1, 2, \dots, n$ , chaque coefficient  $o_k^{(i)}$  est égal au produit matriciel min – max de la  $k$ ème ligne de  $M_i$  par  $I_i$ .

Dans [2], il est décrit comment nous pouvons réduire les matrices  $M_i$  et  $O_i$  de manière à ce qu'elles aient au plus  $\min(\text{card}(D_b), 2^i)$  lignes.

On remarque :

**Proposition 1** *Au moins un coefficient de  $O_n$  est égal à un.*

Ceci peut être vu comme suit. D'après le premier théorème de [2], tout coefficient  $o_k^{(n)}$  de  $O_n$  est de la forme :

$$o_k^{(n)} = \min(\gamma_1, \gamma_2, \dots, \gamma_n), \text{ où } \gamma \in \{\alpha, \beta\}.$$

Puisque nous avons  $\max(\alpha_i, \beta_i) = 1$  pour  $i = 1, 2, \dots, n$ , voir (1), on peut trouver un coefficient de  $O_n$  dont les  $\gamma_1, \gamma_2, \dots, \gamma_n$  sont tous égaux à un.

## 3 Système d'équations pour l'apprentissage

Pour apprendre les paramètres des règles  $s_1, s_2, \dots, s_n$  et  $r_1, r_2, \dots, r_n$  en fonction des données d'apprentissage, nous commençons par introduire un nouveau système d'équations qui est analogue au système  $O_n = M_n \square_{\max}^{\min} I_n$ . Dans ce nouveau système noté  $(\Sigma_n)$ , l'inconnue est un vecteur composé par les paramètres des règles. Ces deux systèmes d'équations sont liés au sens suivant : ils utilisent le même second membre qui décrit la distribution des possibilités de sortie, et les matrices de ces deux systèmes d'équations sont définies par la même construction inductive. En donnant des formules explicites pour calculer les coefficients des matrices impliquées dans les deux systèmes

d'équations, nous établissons une équivalence entre eux.

### 3.1 Définition du systèmes d'équations $(\Sigma_n)$

Nous construisons un système d'équations analogue à  $O_n = M_n \square_{\max}^{\min} I_n$ , qui a la forme suivante :

$$(\Sigma_n) : Y = \Gamma_n \square_{\max}^{\min} X, \quad (2)$$

où :

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{2^n} \end{bmatrix} = O_n \text{ et } X = \begin{bmatrix} s_1 \\ r_1 \\ s_2 \\ r_2 \\ \vdots \\ s_n \\ r_n \end{bmatrix}. \quad (3)$$

La matrice  $\Gamma_n = M_n(\lambda_i, \rho_i)$  est de taille  $(2^n, 2n)$  et est obtenue en remplaçant  $s_i$  par  $\lambda_i$  et  $r_i$  par  $\rho_i$  pour  $i = 1, 2, \dots, n$  dans  $M_n(s_i, r_i)$ .

En prenant  $\Gamma_1 = \begin{bmatrix} \lambda_1 & 1 \\ 1 & \rho_1 \end{bmatrix}$ , la matrice  $\Gamma_i$  pour  $i = 2, 3, \dots, n$  est de taille  $(2^i, 2i)$  et est définie par la construction de matrices par blocs suivante :

$$\Gamma_i = \left[ \begin{array}{c|c} \Gamma_{i-1} & \Lambda_i \\ \hline \Gamma_{i-1} & P_i \end{array} \right], \quad (4)$$

où  $\Lambda_i = \begin{bmatrix} \lambda_i & 1 \\ \lambda_i & 1 \\ \vdots & \vdots \\ \lambda_i & 1 \end{bmatrix}$  et  $P_i = \begin{bmatrix} 1 & \rho_i \\ 1 & \rho_i \\ \vdots & \vdots \\ 1 & \rho_i \end{bmatrix}$  sont de taille  $(2^{i-1}, 2)$ .

Avec les notations de (3), nous établirons (Proposition 3) :

$$O_n = M_n \square_{\max}^{\min} I_n \iff Y = \Gamma_n \square_{\max}^{\min} X. \quad (5)$$

### 3.2 Coefficients de $\Gamma_n$

Pour donner des formules explicites pour les coefficients de  $\Gamma_n$ , nous prenons les notations suivantes.

Fixons  $I = I_n = \{1, 2, \dots, 2^n\}$  et  $J = J_n = \{1, 2, \dots, n\}$ . À chaque élément  $j \in J$ , nous associons une partition de  $I$  :

$$I = \bigcup_{s \in I_{n-j}} I_{s,j}, \quad (6)$$

où  $I_{n-j} = \{1, 2, \dots, 2^{n-j}\}$  et pour chaque  $s \in I_{n-j}$ , on prend  $I_{s,j} = \llbracket (s-1)2^j, s2^j \rrbracket$ .

Nous formons ensuite les partitions de  $I_n$  qui nous permettent d'obtenir les coefficients de  $\Gamma_n$  :

$$I_n = I_j^{n,\top} \cup I_j^{n,\perp} \quad ; \quad j \in J_n \quad (7)$$

où :

$$I_j^{n,\top} = \bigcup_{s \in I_{n-j}} I_{s,j}^\top$$

avec  $I_{s,j}^\top = \llbracket (s-1)2^j, (s-1)2^j + 2^{j-1} \rrbracket$ , et

$$I_j^{n,\perp} = \bigcup_{s \in I_{n-j}} I_{s,j}^\perp$$

avec  $I_{s,j}^\perp = \llbracket (s-1)2^j + 2^{j-1}, s2^j \rrbracket$ .

Nous utiliserons le résultat suivant :

Pour chaque  $1 \leq j \leq n-1$ , la partition  $I_n = I_j^{n,\top} \cup I_j^{n,\perp}$  définie par  $j$  peut-être déduite à partir de la partition de  $I_{n-1} = \{1, \dots, 2^{n-1}\} = I_j^{n-1,\top} \cup I_j^{n-1,\perp}$  par la relation suivante :

$$I_j^{n,\top} = I_j^{n-1,\top} \cup \tau_{2^{n-1}} I_j^{n-1,\top}, \quad (8a)$$

$$I_j^{n,\perp} = I_j^{n-1,\perp} \cup \tau_{2^{n-1}} I_j^{n-1,\perp}, \quad (8b)$$

où  $\tau_{2^{n-1}}$  est l'application  $k \mapsto k + 2^{n-1}$ .

Avec ces notations, on pose :

$$\Gamma_n = M_n(\lambda_i, \rho_i) = [\gamma_{i,j}]_{(i,j) \in I_n \times J_{2n}} \quad (9)$$

**Proposition 2** Pour  $k \in I_n$  et  $j \in J_n$ , les coefficients de  $\Gamma_n$  sont donnés par :

$$\gamma_{k,2j-1} = \begin{cases} \lambda_j & \text{si } k \in I_j^{n,\top} \\ 1 & \text{si } k \in I_j^{n,\perp} \end{cases},$$

$$\gamma_{k,2j} = \begin{cases} 1 & \text{si } k \in I_j^{n,\top} \\ \rho_j & \text{si } k \in I_j^{n,\perp} \end{cases}.$$

**Preuve 1** Ce résultat est prouvé par récurrence sur  $n$  en utilisant (4) et (8).

**Proposition 3** Avec les notations (3), on a :

$$O_n = M_n \square_{\max}^{\min} I_n \iff Y = \Gamma_n \square_{\max}^{\min} X. \quad (11)$$

**Preuve 2** Elle utilise la commutativité de la fonction  $\max$  et la Proposition 2 appliquée aux deux matrices  $\Gamma_n$  et  $M_n$ .

Dans ce qui suit, nous étudions le système  $(\Sigma_n)$ , voir (2), où  $\Gamma_n = M_n(\lambda_i, \rho_i)$  et au moins un coefficient de  $Y$  est égal à un, selon la Proposition 1.

#### 4 Apprentissage des paramètres des règles d'un système à base de règles possibilistes

Le travail fondateur de Sanchez [9] et les résultats de [8] pour la résolution des systèmes d'équations floues de type  $\min - \max$ , seront appliqués pour l'apprentissage des paramètres de règles en fonction des données d'entraînement. Nous rappelons que toute solution du système d'équations  $(\Sigma_n)$ , voir (2), attribue une valeur à chaque paramètre des règles. Dans cette section, nous calculons d'abord la solution minimale potentielle de  $(\Sigma_n)$  et rappelons la condition nécessaire et suffisante de Sanchez pour que  $(\Sigma_n)$  soit compatible. Nous donnons une description complète de l'ensemble des solutions de  $(\Sigma_n)$ , qui entraîne qu'il existe une unique solution maximale. Ces résultats sont énoncés dans le Théorème 1.

##### 4.1 Solution minimale potentielle de $(\Sigma_n)$

Nous utilisons le produit  $\epsilon$  défini par :

$$x\epsilon y = \begin{cases} y & \text{si } x < y \\ 0 & \text{si } x \geq y \end{cases} \quad \text{dans } [0, 1].$$

Nous avons les propriétés utiles suivantes dans  $[0, 1]$  :

$$\max(x\epsilon y_1, x\epsilon y_2) = x\epsilon \max(y_1, y_2), \quad (12a)$$

$$\max(x, x\epsilon y) = \max(x, y). \quad (12b)$$

Nous rappelons que la correspondance  $x \mapsto 1 - x$  définie sur  $[0, 1]$  permute les deux produits matriciels  $\square_{\max}^{\min}$  et  $\square_{\min}^{\max}$ , voir [2]. Il est également facile de vérifier que la même application permute le produit  $\epsilon$  et le produit de Gödel classique :

$$a \rightarrow_G b = \begin{cases} 1 & \text{si } a \leq b \\ b & \text{sinon} \end{cases}.$$

Les systèmes d'équations qui utilisent les produits  $\square_{\max}^{\min}$  ou  $\square_{\min}^{\max}$  sont étudiés dans [8].

Par un résultat fondamental de Sanchez [9], la solution minimale potentielle de  $(\Sigma_n)$  est donnée par le vecteur-colonne suivant :

$$E^{n,\downarrow} = \Gamma_n^t \square_{\epsilon}^{\max} Y = \begin{bmatrix} e_1^{n,\downarrow} \\ e_2^{n,\downarrow} \\ \vdots \\ e_{2n}^{n,\downarrow} \end{bmatrix}. \quad (13)$$

Dans la relation (13), la matrice  $\Gamma_n^t$  est la transposée de la matrice  $\Gamma_n$  et l'opérateur  $\square_{\epsilon}^{\max}$  est le produit matriciel où  $\epsilon$  est le produit et  $\max$  l'addition.

**Proposition 4** Pour chaque  $j = 1, 2, \dots, n$ , les coefficients de  $E^{n,\downarrow}$  sont donnés par :

$$e_{2j-1}^{n,\downarrow} = \lambda_j \epsilon y_j^{n,\top} \quad \text{et} \quad e_{2j}^{n,\downarrow} = \rho_j \epsilon y_j^{n,\perp}, \quad (14)$$

où :

$$y_j^{n,\top} = \max_{i \in I_j^{n,\top}} y_i \quad \text{et} \quad y_j^{n,\perp} = \max_{i \in I_j^{n,\perp}} y_i. \quad (15)$$

**Preuve 3** Elle s'obtient par un calcul direct qui utilise (12a) et la Proposition 2.

Dans ce qui suit, le vecteur-colonne  $E^{n,\downarrow}$  sera utilisé pour rappeler la condition nécessaire et suffisante de Sanchez [9] pour la compatibilité de  $(\Sigma_n)$ .

## 4.2 Condition nécessaire et suffisante pour la compatibilité de $(\Sigma_n)$

Pour caractériser la compatibilité de  $(\Sigma_n)$ , nous introduisons le vecteur-colonne suivant :

$$\tilde{Y} = \Gamma_n \square_{\max}^{\min} E^{n,\downarrow} = [\tilde{y}_i]_{1 \leq i \leq 2^n}. \quad (16)$$

Par les travaux de Sanchez [9], le système  $(\Sigma_n)$  est compatible si et seulement si :

$$Y = \tilde{Y}. \quad (17)$$

**Proposition 5** Pour  $i = 1, 2, \dots, 2^n$ , le coefficient  $\tilde{y}_i$  de  $\tilde{Y}$  est donné par :

$$\tilde{y}_i = \min_{1 \leq j \leq n} z_{ij}; z_{ij} = \begin{cases} \max(\lambda_j, y_j^{n,\top}) & \text{si } i \in I_j^{n,\top} \\ \max(\rho_j, y_j^{n,\perp}) & \text{si } i \in I_j^{n,\perp} \end{cases}. \quad (18)$$

où  $y_j^{n,\top}$  et  $y_j^{n,\perp}$  sont définis dans (15).

**Preuve 4** Elle s'obtient par un calcul direct qui utilise (12b) et Proposition 4.

On déduit de cette proposition que  $\tilde{Y} \geq Y$ , ce qui signifie :

$$\forall i \in \{1, 2, \dots, 2^n\}, \tilde{y}_i \geq y_i. \quad (19)$$

Nous énonçons maintenant les principaux résultats pour la résolution du système  $(\Sigma_n)$  :

**Théorème 1** Si le système  $(\Sigma_n)$  est compatible, on a :

- i) L'unique solution minimale est le vecteur-colonne  $E^{n,\downarrow}$  de (13).
- ii) Il n'y a qu'une seule solution maximale :

$$E^{n,\uparrow} = \begin{bmatrix} e_1^{n,\uparrow} \\ e_2^{n,\uparrow} \\ \vdots \\ e_{2^n}^{n,\uparrow} \end{bmatrix}. \quad (20)$$

où pour  $j = 1, 2, \dots, n$  :

$$e_{2^j-1}^{n,\uparrow} = y_j^{n,\top} \text{ and } e_{2^j}^{n,\uparrow} = y_j^{n,\perp}. \quad (21)$$

iii) Un vecteur colonne  $X = [x_k]_{1 \leq k \leq 2^n}$  est une solution du système  $(\Sigma_n)$  si et seulement si :

$$E^{n,\downarrow} \leq X \leq E^{n,\uparrow}, \quad (22)$$

ce qui signifie :

$$\forall k \in \{1, 2, \dots, 2^n\}, e_k^{n,\downarrow} \leq x_k \leq e_k^{n,\uparrow}.$$

Nous rappelons que si le système  $(\Sigma_n)$  est compatible, Sanchez a montré que  $E^{n,\downarrow}$  est son unique solution minimale [8, 9].

La preuve du Théorème 1 se compose de deux parties. Dans une première partie, nous comparons la compatibilité du système  $(\Sigma_n)$  et celle du système à  $2^{n-1}$  équations suivant :

$$(S_{n-1}) : \Gamma_{n-1} \square_{\max}^{\min} X' = B,$$

où les coefficients de son second membre  $B = [b_i]_{1 \leq i \leq 2^{n-1}}$  sont définis à partir de ceux du second membre  $Y$  de  $(\Sigma_n)$  :

$$b_i = \max(y_i, y_{i+2^{n-1}}).$$

Nous établissons un lien entre la compatibilité et l'ensemble des solutions des deux systèmes  $(\Sigma_n)$  et  $(S_{n-1})$ , ce qui nous permet de ramener la résolution de  $(\Sigma_n)$  à celle de  $(S_{n-1})$ .

Dans une seconde partie, nous itérons la construction  $(\Sigma_n) \rightsquigarrow (S_{n-1})$ , ce qui permet d'obtenir une suite  $(\Sigma_k)_{1 \leq k \leq n}$  de systèmes dont le terme en  $k = n$  est le système initial. La preuve du Théorème 1 pour le système  $(\Sigma_1)$  est facile car la matrice de ce système comporte deux lignes et deux colonnes. En se basant sur les résultats de la première partie et un raisonnement par récurrence, nous établissons les énoncés du Théorème 1.

## 5 Exemple

Nous considérons un système composé de  $n = 3$  règles possibilistes parallèles et de la donnée

d'entrainement suivante :

$$\begin{bmatrix} \lambda_1 \\ \rho_1 \\ \lambda_2 \\ \rho_2 \\ \lambda_3 \\ \rho_3 \end{bmatrix} = \begin{bmatrix} 0.1 \\ 1 \\ 1 \\ 0.8 \\ 1 \\ 0.3 \end{bmatrix} \text{ et } Y = \begin{bmatrix} 0.3 \\ 1 \\ 0.3 \\ 0.8 \\ 0.3 \\ 0.7 \\ 0.3 \\ 0.7 \end{bmatrix}.$$

Le système  $(\Sigma_3) : \Gamma_3 \square_{\max}^{\min} X$  est :

$$Y = \begin{bmatrix} 0.1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0.1 & 1 & 1 & 0.8 & 1 & 1 \\ 1 & 1 & 1 & 0.8 & 1 & 1 \\ 0.1 & 1 & 1 & 1 & 1 & 0.3 \\ 1 & 1 & 1 & 1 & 1 & 0.3 \\ 0.1 & 1 & 1 & 0.8 & 1 & 0.3 \\ 1 & 1 & 1 & 0.8 & 1 & 0.3 \end{bmatrix} \square_{\max}^{\min} \begin{bmatrix} s_1 \\ r_1 \\ s_2 \\ r_2 \\ s_3 \\ r_3 \end{bmatrix}$$

et le système  $(\Sigma_2) = (S_2) = B = \Gamma_2 \square_{\max}^{\min} X'$  :

$$\begin{bmatrix} 0.3 \\ 1 \\ 0.3 \\ 0.8 \end{bmatrix} = B = \begin{bmatrix} 0.1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0.1 & 1 & 1 & 0.8 \\ 1 & 1 & 1 & 0.8 \end{bmatrix} \square_{\max}^{\min} \begin{bmatrix} s_1 \\ r_1 \\ s_2 \\ r_2 \end{bmatrix}.$$

Le calcul de la solution minimale potentielle de chaque système s'obtient par :

$$\begin{aligned} y_1^{3,\top} = b_1^{2,\top} = 0.3, & \quad y_1^{3,\perp} = b_1^{2,\perp} = 1, \\ y_2^{3,\top} = b_2^{2,\top} = 1, & \quad y_2^{3,\perp} = b_2^{2,\perp} = 0.8, \\ y_3^{3,\top} = 1, & \quad y_3^{3,\perp} = 0.7. \end{aligned}$$

Donc :

$$E^{3,\downarrow} = \begin{bmatrix} e_1^{3,\downarrow} \\ e_2^{3,\downarrow} \\ e_3^{3,\downarrow} \\ e_4^{3,\downarrow} \\ e_5^{3,\downarrow} \\ e_6^{3,\downarrow} \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0.7 \end{bmatrix}$$

et

$$E^{2,\downarrow} = \begin{bmatrix} e_1^{2,\downarrow} \\ e_2^{2,\downarrow} \\ e_3^{2,\downarrow} \\ e_4^{2,\downarrow} \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

$$\text{où } e_1^{3,\downarrow} = e_1^{2,\downarrow} = \lambda_1 \epsilon y_1^{3,\top}, e_2^{3,\downarrow} = e_2^{2,\downarrow} = \rho_1 \epsilon y_1^{3,\perp}, \\ e_3^{3,\downarrow} = e_3^{2,\downarrow} = \lambda_2 \epsilon y_2^{3,\top}, e_4^{3,\downarrow} = e_4^{2,\downarrow} = \rho_2 \epsilon y_2^{3,\perp}, \\ e_5^{3,\downarrow} = \lambda_3 \epsilon y_3^{3,\top} \text{ et } e_6^{3,\downarrow} = \rho_3 \epsilon y_3^{3,\perp}.$$

Le système  $(\Sigma_3)$  est compatible : par exemple on peut facilement vérifier que nous avons  $\tilde{Y} = \Gamma_3 \square_{\max}^{\min} E^{3,\downarrow} = Y$ .

La solution maximale de chaque système est :

$$E^{3,\uparrow} = \begin{bmatrix} e_1^{3,\uparrow} \\ e_2^{3,\uparrow} \\ e_3^{3,\uparrow} \\ e_4^{3,\uparrow} \\ e_5^{3,\uparrow} \\ e_6^{3,\uparrow} \end{bmatrix} = \begin{bmatrix} 0.3 \\ 1 \\ 1 \\ 0.8 \\ 1 \\ 0.7 \end{bmatrix}$$

et

$$E^{2,\uparrow} = \begin{bmatrix} e_1^{2,\uparrow} \\ e_2^{2,\uparrow} \\ e_3^{2,\uparrow} \\ e_4^{2,\uparrow} \end{bmatrix} = \begin{bmatrix} 0.3 \\ 1 \\ 1 \\ 0.8 \end{bmatrix}$$

$$\text{où } e_1^{3,\uparrow} = e_1^{2,\uparrow} = y_1^{3,\top}, e_2^{3,\uparrow} = e_2^{2,\uparrow} = y_1^{3,\perp}, e_3^{3,\uparrow} = e_3^{2,\uparrow} = y_2^{3,\top}, \\ e_4^{3,\uparrow} = e_4^{2,\uparrow} = y_2^{3,\perp}, e_5^{3,\uparrow} = y_3^{3,\top} \text{ et } e_6^{3,\uparrow} = y_3^{3,\perp}.$$

À partir du Théorème 1, on déduit que  $X =$

$$\begin{bmatrix} s_1 \\ r_1 \\ s_2 \\ r_2 \\ s_3 \\ r_3 \end{bmatrix}$$

est une solution si et seulement si :

$$\begin{bmatrix} 0.3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0.7 \end{bmatrix} \leq \begin{bmatrix} s_1 \\ r_1 \\ s_2 \\ r_2 \\ s_3 \\ r_3 \end{bmatrix} \leq \begin{bmatrix} 0.3 \\ 1 \\ 1 \\ 0.8 \\ 1 \\ 0.7 \end{bmatrix}$$

i.e.,  $s_1 = 0.3, 0 \leq r_2 \leq 0.8$  et  $r_3 = 0.7$ .

## 6 Conclusion

Dans cet article, nous avons introduit un paradigme d'apprentissage des paramètres des règles d'un système à base de règles possibilistes, en résolvant un système d'équations analogue à celui de Farreny et Prade. L'approche

proposée s'appuie sur les méthodes et résultats de résolution de systèmes d'équations floues de type min-max. Pour un système d'équations compatible, nous avons déterminé l'ensemble complet de ses solutions et déduit qu'il a une solution maximale unique. En tenant compte du résultat de Sanchez, si notre système d'équations est compatible, alors ses solutions extrêmes sont uniques. Pour notre système d'équations, comme dans [2], il est possible de réduire le nombre de lignes de la matrice  $\Gamma_n$  (et  $Y$ ). Comme perspectives pour l'apprentissage, nous pouvons trouver des moyens de fusionner des ensembles de solutions provenant de plusieurs données d'entraînement. Dans le cas d'un système non compatible associé à une donnée d'entraînement, nous pouvons également rechercher la meilleure approximation de la distribution des possibilités de sortie ciblée dont le système correspondant est compatible.

## Références

- [1] Baaj, I. 2022, *Explainability of possibilistic and fuzzy rule-based systems*, thèse de doctorat, Sorbonne Université.
- [2] Baaj, I., J.-P. Poli, W. Ouerdane et N. Maudet. 2021, «Inférence min-max pour un système à base de règles possibilistes», dans *Rencontres francophones sur la logique floue et ses applications*, Sorbonne Université, Cepadues, Paris, France, p. 233–240. I.S.B.N. : 9782364939066.
- [3] Dubois, D. et H. Prade. 1988, *Possibility theory : an approach to computerized processing of uncertainty*, Plenum Press, New York, ISBN 0306425203.
- [4] Dubois, D. et H. Prade. 2015, «Possibility theory and its applications : Where do we stand?», dans *Springer Handbook of Computational Intelligence*, édité par J. Kacprzyk et W. Pedrycz, Springer Berlin Heidelberg, ISBN 978-3-662-43505-2, p. 31–60.
- [5] Dubois, D. et H. Prade. 2020, «From possibilistic rule-based systems to machine learning - a discussion paper», dans *Scalable Uncertainty Management*, édité par J. Davis et K. Tabia, Springer International Publishing, Cham, ISBN 978-3-030-58449-8, p. 35–51.
- [6] Farreny, H. et H. Prade. 1986, «Default and inexact reasoning with possibility degrees», *IEEE transactions on systems, man, and cybernetics*, vol. 16, n° 2, p. 270–276.
- [7] Farreny, H. et H. Prade. 1990, «Explications de raisonnements dans l'incertain», *Revue d'intelligence artificielle*, vol. 4, n° 2, p. 43–75.
- [8] Peeva, K. 2013, «Resolution of fuzzy relational equations—method, algorithm and software with applications», *Information Sciences*, vol. 234, p. 44–63.
- [9] Sanchez, E. 1976, «Resolution of composite fuzzy relation equations», *Information and control*, vol. 30, n° 1, p. 38–48.