

Aprendizaje de Pesos y Selección Evolutiva para la Reducción de la Base de Reglas Difusas*

Jesús Alcalá¹, Rafael Alcalá¹, Oscar Cordon¹, Francisco Herrera¹

Resumen— En problemas multidimensionales complejos con una elevada no linealidad en la relación entrada-salida, pueden encontrarse reglas inconsistentes y redundantes en la base de reglas del modelo difuso lingüístico, lo que repercute en una pérdida de precisión e interpretabilidad. Además, las reglas podrían no cooperar todo lo bien que se desearía.

Es sabido que el uso de reglas ponderadas como refinamiento local de reglas lingüísticas, permite al modelo difuso lingüístico hacer frente a reglas ineficientes mejorando la robustez, flexibilidad y capacidad de modelado del sistema. Por otro lado, la selección de reglas realiza una reducción de una base de reglas difusas previamente obtenida, eliminando las reglas ineficientes y/o redundantes con el fin de mejorar la cooperación entre ellas, reduciendo la complejidad del modelo y mejorando la legibilidad del sistema. Un sistema compacto con pocas reglas requiere un menor esfuerzo para ser interpretado. Ya que estos dos enfoques presentan características complementarias, podrían ser combinados mediante la evolución conjunta de pesos y selección de reglas para la obtención de modelos difusos lingüísticos compactos, más precisos y que mantengan un buen grado de interpretabilidad.

Palabras clave— Sistemas basados en reglas difusas, modelado difuso lingüístico, selección de reglas, reglas difusas ponderadas, algoritmos genéticos.

I. INTRODUCCIÓN

UNO de los problemas asociados con el Modelado Difuso Lingüístico es carecer de precisión cuando se trata de modelar problemas complejos. Esto es debido a la inflexibilidad del concepto de variable lingüística, el cual impone fuertes restricciones a la estructura de las reglas difusas. En este tipo de modelado, la precisión y la interpretabilidad son propiedades contradictorias y directamente dependientes del proceso de aprendizaje y/o de la estructura del modelo.

En los problemas multidimensionales complejos con una elevada no linealidad en la relación entrada-salida, es usual encontrar en la base

muchas reglas redundantes, inconsistentes y conflictivas, lo que empeora el funcionamiento y la interpretabilidad del modelo difuso lingüístico. En este caso, las reglas no cooperan apropiadamente.

Para solucionar estos problemas, se han considerado muchas y diferentes mejoras al Modelado Difuso Lingüístico en la literatura [1]. Todos estos enfoques tienen la idea común de mejorar el razonamiento interpolativo del modelo difuso lingüístico, inducido por una mejor cooperación entre las reglas del modelo aprendido. La cooperación entre las reglas puede ser mejorada actuando sobre tres componentes diferentes del modelo: la base de datos, la base de reglas y la base de conocimiento. Si nos centramos en la base de reglas, podemos mejorar la cooperación actuando sobre estos componentes:

- *Selección o reducción de reglas* [2], [3], [4], [5], [6], [7], [8], [9]: Implica la obtención de un subconjunto de reglas optimizado a partir de una base de reglas previa.
- *Uso de reglas con varios consecuentes* [10], [11]: Este enfoque permite que la base de reglas contenga reglas donde cada combinación de consecuentes presente dos o más consecuentes asociados cuando sea necesario.
- *Uso de reglas lingüísticas ponderadas* [12], [13], [14]: En este caso, se considera un parámetro adicional para cada regla que representa su grado de importancia en el proceso de inferencia, en lugar de considerar todas las reglas con la misma importancia como suele ser habitual.
- *Inducción directa de cooperación entre reglas* [15], [16]: Este tipo de métodos se basan en la mejora del proceso de generación de reglas mediante el uso de un criterio global de rendimiento que considera conjuntamente la acción de las distintas reglas.

El aprendizaje de pesos para las reglas lingüísticas y la selección de reglas presentan características complementarias. Por un lado, es sabido que el uso de reglas ponderadas como refinamiento local de reglas lingüísticas, permite al modelo difuso lingüístico hacer frente a reglas in-

* Financiado por el Ministerio de Ciencia y Tecnología y los Fondos FEDER bajo el proyecto TIC-2002-04036-C05-01.

¹ Departamento de Ciencias de la Computación e I.A., Universidad de Granada, E-18071 Granada, España. E-mail: {jalcala, alcalá, ocordon, herrera}@decsai.ugr.es

eficientes y redundantes mejorando la robustez, flexibilidad y capacidad de modelado del sistema. Por otro lado, la selección de reglas realiza una simplificación de una base de reglas difusas previamente obtenida, eliminando las reglas ineficientes y redundantes con el fin de mejorar la cooperación entre ellas. Además, reduciendo la complejidad del modelo se puede mejorar la legibilidad del sistema, es decir, un sistema compacto con pocas reglas requiere un menor esfuerzo para ser interpretado. Puesto que estos dos enfoques tienen características complementarias, son susceptibles de ser combinados.

En este trabajo analizamos la hibridación de ambas técnicas con el objetivo de obtener modelos difusos lingüísticos más simples y precisos. Para seleccionar las reglas que mejor cooperan y los pesos asociados a ellas utilizamos un Algoritmo Genético (AG) [17], [18].

En esta propuesta se plantea el uso combinado de reglas difusas ponderadas y de selección de reglas. *Esto puede ser entendido como un meta-método sobre algún otro método de generación de reglas lingüísticas, desarrollado para obtener modelos difusos lingüísticos más simples y precisos al seleccionar solo las reglas con una buena cooperación.* De la combinación de esta técnica con diferentes métodos de aprendizaje de reglas difusas, surgen distintos mecanismos de aprendizaje. Nosotros consideraremos como métodos para la generación del conjunto de reglas inicial el método de Wang y Mendel [19] y una extensión de este método para obtener reglas difusas con doble consecuente [10].

Esta contribución se organiza de la siguiente manera. En las secciones II y III se analiza el uso de reglas ponderadas y la selección de reglas, considerándolas como dos formas complementarias de mejorar el rendimiento del modelo lingüístico. En las secciones IV y V presentamos, respectivamente, la estrategia de aprendizaje propuesta y el funcionamiento del proceso de optimización evolutivo para la selección de reglas y el aprendizaje de pesos. En la sección VI se muestran los resultados que hemos obtenido en un problema real. Finalmente, en la sección VII apuntamos algunas conclusiones finales.

II. REGLAS LINGÜÍSTICAS PONDERADAS

Usualmente, las reglas difusas ponderadas [12], [14], [20] han sido consideradas para mejorar la forma en la cual dichas reglas interactúan, aumentando la precisión del modelo así obtenido. De esta manera, el uso de pesos supone una extensión efectiva del sistema de razonamiento difuso convencional que permite el ajuste del sistema

a nivel de regla [12], [14].

Cuando los pesos se aplican a reglas completas, cada peso modula el grado de activación de la correspondiente regla en el proceso de cálculo de la salida del sistema. Para el entendimiento humano, esto es muy cercano a considerar el peso de cada regla como un grado de importancia asociado a la misma, determinando como dicha regla interactúa con sus vecinas. Además, sólo se considerarán pesos con valores en $[0, 1]$ puesto que ello preserva la legibilidad del modelo. Desde este punto de vista, el uso de pesos representa un campo de trabajo ideal para el modelado lingüístico cuando se pretende un compromiso entre precisión e interpretabilidad.

Para ello, utilizamos la estructura de regla y el sistema de inferencia propuestos en [14]:

SI X_1 es A_1 y ... y X_n es A_n
ENTONCES Y es B con $[w]$,

donde X_i (Y) son las variables de entrada (salida), A_i (B) son las etiquetas lingüísticas utilizadas en la regla, w es un valor real representando el peso de la regla, y *con* es un operador que modela la aplicación del peso asignado a dicha regla.

Con esta estructura, el razonamiento difuso debe ser extendido. El enfoque clásico es inferir utilizando una estrategia de defuzzificación que trabaje en modo B [21] y calcular la salida real como la siguiente *suma ponderada*:

$$y_0 = \frac{\sum_i m_i \cdot w_i \cdot P_i}{\sum_i m_i \cdot w_i},$$

donde y_0 es el valor real obtenido tras el proceso de defuzzificación, m_i es el grado de emparejamiento de la i -ésima regla, w_i es el peso asociado a dicha regla, y P_i es el valor característico del conjunto difuso de salida obtenido a partir de la misma. En este trabajo, consideramos el *centro de gravedad* como valor característico y la *t-norma del mínimo* como operador de implicación y conjunción.

Una aproximación simple para el aprendizaje de reglas ponderadas consistiría en utilizar un método de optimización, por ejemplo AGs, para obtener los pesos de un conjunto de reglas previamente obtenido.

III. SELECCIÓN DE REGLAS COOPERATIVAS

En problemas multidimensionales complejos con una elevada no linealidad en la relación entrada-salida, es usual encontrar reglas inconsistentes y redundantes en la base de reglas (especialmente cuando estas son generadas considerando solamente conocimiento experto). Por otro lado, en problemas de gran dimensión el número

de reglas de la base crece exponencialmente a medida que se añaden entradas. Un conjunto de reglas grandes puede contener muchas reglas redundantes, inconsistentes y conflictivas. Este tipo de reglas empeoran el funcionamiento y la interpretabilidad del modelo.

Los métodos de selección de reglas seleccionan un subconjunto de éstas de un conjunto de reglas previamente obtenido. De esta manera conseguimos eliminar reglas inconsistentes y redundantes, mejorando la cooperación entre las reglas difusas y provocando una mejora potencial en el funcionamiento del sistema. Además, reducimos la complejidad del modelo, mejorando así su interpretabilidad.

Los métodos de reducción han sido formulados como Redes Neuronales, técnicas de agrupamiento y métodos de transformación ortogonal, y algoritmos basados en medidas de similitud entre otros [2], [3], [6], [7], [8], [9]. En [22] se propuso una aproximación diferente, la cual intentaba reducir el tamaño de la base de reglas transformando las reglas difusas elementales a forma DNF.

El uso de un AG para encontrar el subconjunto óptimo de reglas difusas está motivado en las siguientes situaciones:

- la integración de un conjunto de reglas expertas y un conjunto de reglas difusas extraídas mediante métodos de aprendizaje automático [23].
- la selección un conjunto de reglas cooperativas a partir de un conjunto de reglas candidatas [24], [10], [25], [4], [26], [5].
- la selección de reglas de una base de conocimiento previamente definida junto con la selección de las etiquetas para las variables consecuente [27].
- la selección de reglas junto con un refinamiento de las funciones de pertenencia, codificando todo (reglas y parámetros) en un solo cromosoma [28].
- el desarrollo de modelos difusos compactos mediante la reducción de la complejidad, para lo cual se combinan: agrupamiento difuso, reducción de reglas por técnicas ortogonales, simplificación dirigida por similitudes y optimización genética [29].

En nuestro caso son particularmente interesante la segunda y la cuarta. Nosotros proponemos seleccionar un conjunto de reglas cooperativas a partir de un conjunto de reglas candidatas junto con el aprendizaje de pesos para ellas, codificando todo en un solo cromosoma. El objetivo que perseguimos con ello es:

- Mejorar la precisión del modelo lingüístico, seleccionando el conjunto de reglas que mejor co-

operan.

- Obtener modelos lingüísticos simples que mantengan un buen grado de interpretabilidad.

IV. COMBINACIÓN DEL APRENDIZAJE DE PESOS Y SELECCIÓN DE REGLAS

Como ya dijimos, la combinación del aprendizaje de los pesos asociados a las reglas y del proceso de selección de reglas prodría generar modelos lingüísticos más simples y precisos, siendo más fáciles de comprender y manteniendo un nivel de interpretabilidad adecuado.

Para generar los modelos lingüísticos combinando ambos enfoques vamos a realizar dos pasos:

1. Generar un conjunto de reglas difusas mediante un proceso de generación específico.
2. El segundo paso consta de dos tareas:
 - Selección genética de un subconjunto de reglas que presentan buena cooperación.
 - Aprendizaje genético de los pesos asociados a las reglas seleccionadas.

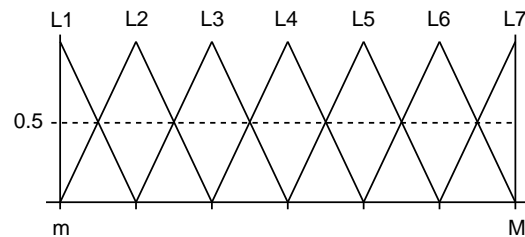


Fig. 1. Representación gráfica de la partición difusa.

Vamos a considerar particiones simétricas y funciones de pertenencia triangulares (ver Figura 1) para generar el conjunto de reglas difusas candidatas. Con este objetivo hemos utilizado dos métodos diferentes pero similares, dependiendo de la estructura de regla que queremos que tenga el modelo obtenido:

- *Basado en la estructura clásica de las reglas.* Genera un conjunto de reglas difusas con la estructura usual. El proceso considerado es el propuesto por Wang y Mendel [19].
- *Basado en la estructura de reglas con doble consecuente.* Para ello, tomamos el primer paso de la metodología de Modelado Lingüístico Preciso (ALM) [10] y el proceso propuesto por Wang y Mendel [19]. Este proceso conlleva dividir el espacio de entrada y salida en regiones difusas diferentes, generando la mejor regla que cubra cada ejemplo y seleccionando después, para cada

espacio de entrada, las dos reglas que mejor grado de cubrimiento tengan. La estructura de las reglas con doble consecuente es

Si X_1 es A_1 y \dots y X_n es A_n
ENTONCES Y es $\{B_1, B_2\}$,

donde X_i (Y) son las variables de entrada (salida), A_i es la etiqueta lingüística de la variable de entrada i -ésima y B_1 y B_2 son las etiquetas lingüísticas de la variable de salida.

Tanto el proceso de aprendizaje de pesos como el de selección de reglas actúan sobre las reglas con estructura clásica (consecuente único), por lo que la base de reglas difusas generada por esta aproximación sufrirá un preprocesamiento que dividirá cada regla en dos reglas simples, cada una de ellas con una de las etiquetas de la variables de salida.

Para seleccionar el subconjunto de reglas que mejor cooperan y asociarles pesos vamos a considerar un AG que codifique todo (reglas y pesos) en un cromosoma. Este algoritmo lo presentaremos a continuación.

V. ALGORITMO GENÉTICO PARA APRENDIZAJE DE PESOS Y SELECCIÓN DE REGLAS

Hemos utilizado un AG para realizar la selección de reglas junto con el aprendizaje de pesos para las reglas. En este algoritmo hemos considerado un esquema generacional [18] clásico elitista (mantiene en la población actual el mejor individuo de la población anterior) junto con el *procedimiento de muestreo universal estocástico de Baker* [30].

La **probabilidad de selección** de los cromosomas se calcula siguiendo un ranking lineal [30], ordenando los cromosomas según su fitness y asignado a cada uno de ellos la siguiente probabilidad ($p_s(C^p)$):

$$p_s(C^p) = \frac{1}{N_C} \cdot (\eta_{max} - (\eta_{max} - \eta_{min}) \cdot \frac{rank(C^p) - 1}{N_C - 1})$$

donde N_C es el número de cromosomas y $\eta_{min} \in [0, 1]$ especifica del número de copias esperadas del peor cromosoma (se esperan $\eta_{max} = 2 - \eta_{min}$ copias del mejor).

Se utiliza un **doble esquema de codificación** ($C = C_1 + C_2$) para la selección de reglas y el aprendizaje de pesos:

- Para la parte C_1 , el esquema de codificación genera cadenas con codificación binaria de longitud m (número de reglas simples en el conjunto de

reglas previamente generado). Dependiendo de si una regla está seleccionada o no, el gen correspondiente tendrá asignado 1 ó 0 respectivamente.

- Para la parte C_2 , el esquema de codificación genera cadenas con codificación real de longitud m . El valor de cada gen indica el peso utilizado para la regla correspondiente. Cada gen puede toma un valor dentro del intervalo $[0, 1]$.

Así, el cromosoma C^p es codificado de la siguiente manera:

$$\begin{aligned} C_1^p &= (c_{11}^p, \dots, c_{1m}^p) \mid c_{1i}^p \in \{0, 1\}, \\ C_2^p &= (c_{21}^p, \dots, c_{2m}^p) \mid c_{2i}^p \in [0, 1], \\ C^p &= C_1^p C_2^p. \end{aligned}$$

La **población inicial** se obtiene incluyendo un individuo C^1 que tiene todos los genes con valor 1 (en ambas partes) y el resto de individuos son generados aleatoriamente.

$$\forall k \in \{1, \dots, m\}, c_{1k}^1 = 1 \text{ y } c_{2k}^1 = 1, 0.$$

Para **evaluar** el cromosoma p -ésimo (C^p) vamos a usar el *error cuadrático medio* (ECM):

$$ECM(E, BR(C^p)) = \frac{\sum_{e^l \in E} (ey^l - s(ex^l))^2}{2 \cdot |E|},$$

donde E es el conjunto de entrenamiento y $s(ex^l)$ es el valor de salida obtenido desde la base de reglas codificada en el cromosoma C^p ($BR(C^p)$) cuando la entrada es $ex^l = (ex_1^l, \dots, ex_n^l)$ y ey^l es la salida deseada.

El **operador de cruce** dependerá de la parte del cromosoma sobre el que se aplica: en la parte C_1 se utiliza el *operador de cruce en 2 puntos* mientras que en la parte C_2 se considera el *operador de cruce max-min-aritmético* [31].

El operador de cruce en 2 puntos intercambia los fragmentos de los padres que se encuentran entre dos puntos seleccionados aleatoriamente, dando como resultado dos descendientes. Cuando usamos el operador de cruce max-min-aritmético sobre la segunda parte, si $C_2^v = (c_{21}^v, \dots, c_{2k}^v, \dots, c_{2m}^v)$ y $C_2^w = (c_{21}^w, \dots, c_{2k}^w, \dots, c_{2m}^w)$ van a ser cruzados, los descendientes resultantes son los dos mejores de los cuatro siguientes descendientes:

$$O_2^1 = aC_2^w + (1 - a)C_2^v,$$

$$O_2^2 = aC_2^v + (1 - a)C_2^w,$$

$$O_2^3 \text{ con } c_{2k}^3 = \min\{c_{2k}^v, c_{2k}^w\},$$

$$O_2^4 \text{ con } c_{2k}^4 = \max\{c_{2k}^v, c_{2k}^w\},$$

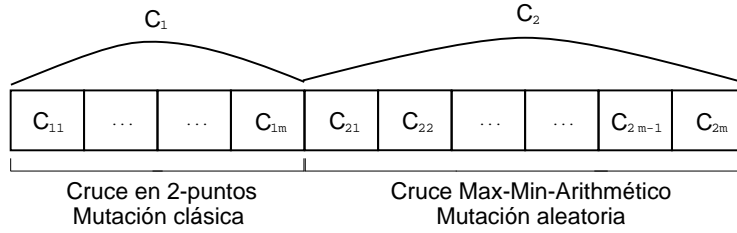


Fig. 2. Representación genética y alcance de aplicación de los operadores.

TABLA I
MÉTODOS CONSIDERADOS PARA SER COMPARADOS.

Ref.	Métodos	Description
[19]	WM	Método ad hoc muy conocido para obtener reglas simples
[10]	DC	Método para obtener reglas con doble consecuente (primer paso de ALM)
[10], [23]	S	AG para la selección de reglas (segundo paso de ALM o la parte C_1 de WS)
---	W	AG para la generación de pesos (parte C_2 de WS)
---	WS	AG propuesto para generación de pesos y selección

De esta manera generamos ocho descendientes al combinar los dos descendientes obtenidos del cruce en la parte C_1 con los cuatro que se obtienen del cruce en la parte C_2 . Los dos mejores descendientes reemplazarán a los dos padres correspondientes en la población.

El **operador de mutación** consiste en asignar 0 ó 1 si el valor del gen es 1 ó 0, respectivamente, y el gen se encuentra en la parte C_1 , y consiste en tomar un valor aleatorio dentro del intervalo $[0, 1]$ si el gen correspondiente está en la parte C_2 .

La Figura 2 representa el ámbito de aplicación de estos operadores.

VI. EXPERIMENTOS

Con objeto de analizar el comportamiento de los métodos generados hemos escogido un problema real, la estimación de la longitud de las líneas de baja tensión de una compañía eléctrica [32].

Hemos utilizado dos métodos diferentes para generar el conjunto inicial de reglas candidatas que van a ser ponderadas y/o seleccionadas (ver sección IV): el método de Wang y Mendel (WM) [19] y una extensión de este método para obtener reglas difusas con doble consecuente (DC) basado en la metodología ALM [10]. Con el objetivo de ver las ventajas de la acción combinada de las reglas ponderadas y de la selección de reglas, hemos realizado tres estudios diferentes para ambos enfoques:

- Solo considerando selección de reglas (S).
- Solo considerando reglas ponderadas (W).
- Considerando reglas ponderadas y selección de reglas a la vez (WS) —algoritmo propuesto en este trabajo—.

La Tabla I presenta una breve descripción de los métodos estudiados.

Consideramos el *centro de gravedad ponderado por el grado de emparejamiento* [21] como operador de defuzzificación y la *t-norma mínimo* como operador de implicación y conjunción.

Los valores de los parámetros utilizados en todos los experimentos presentados son: 61 individuos en la población, 1,000 generaciones, 0.6 como probabilidad de cruce, 0.2 como probabilidad de mutación por cromosoma, 0.35 para el factor a del operador de cruce max-min-aritmético y 0.75 para el parámetro η_{min} de la probabilidad de selección.

A. Estimación de la longitud de las líneas de baja tensión de una compañía eléctrica.

En muchas ocasiones, medir la cantidad de línea eléctrica empleada por una compañía puede ser de gran utilidad por varias razones como, por ejemplo, estimar los costes de mantenimiento de la red de baja tensión, que fue el principal objetivo planteado en la provincia de Asturias. Este

TABLA II
RESULTADOS OBTENIDOS PARA EL PROBLEMA DE LA LONGITUD DE LAS LÍNEAS DE BAJA TENSIÓN

Método	#R \leftarrow (SC+DC)	ECM $_{ent}$	ECM $_{test}$
WM	24 --	222,654	239,962
WM-S	17 --	214,177	265,179
WM-W	24 --	191,577	221,583
WM-WS	20 --	191,565	219,370
Considerando reglas con doble consecuente			
DC	24 (14+10)	231,132	259,973
DC-S (ALM)	17 (14+ 3)	155,898	178,534
DC-W	24 (14+10)	144,983	191,053
DC-WS	18 (15+ 3)	144,656	177,897

SC = Simple Consecuente, DC = Doble Consecuente.

tipo de medición debe ser llevado a cabo por técnicas indirectas ya que, al estar las redes tendidas en pequeños núcleos rurales, su medición directa resulta muy cara. Así, abordar el problema con modelado difuso lingüístico implica encontrar un modelo que relacione la longitud total de línea de baja tensión instalada en un núcleo rural, con el número de habitantes y su radio [32].

Para ello, dividimos aleatoriamente al 80-20% la muestra de datos de 495 núcleos rurales en dos conjuntos de entrenamiento y test, formados por tanto por 396 y 99 datos respectivamente. Las particiones lingüísticas consideradas se componen de *siete términos lingüísticos* con forma triangular (véase la figura 1). Las etiquetas correspondientes, $\{L_1, L_2, L_3, L_4, L_5, L_6, L_7\}$, representan:

- L_1 : Extremadamente pequeño.
- L_2 : Muy pequeño.
- L_3 : Pequeño.
- L_4 : Medio.
- L_5 : Grande.
- L_6 : Muy grande.
- L_7 : Extremadamente grande.

B. Resultados y análisis.

Los resultados obtenidos para los cuatro métodos analizados se muestran en la Tabla II, donde #R representa el número de reglas, y ECM $_{ent}$ y ECM $_{test}$ representan el error obtenido sobre el conjunto de entrenamiento y de prueba respectivamente. En cada tabla se muestran los mejores resultados obtenidos. La experimentación fue realizada en un PENTIUM III con una frecuencia de reloj de 550 MHz y 128 MB de memoria RAM. El tiempo de ejecución para los distintos algoritmos no excedió los 20 minutos.

Centrándonos en la aproximación WM, el mod-

elo WM-WS es el que tiene un mejor funcionamiento, mejorando en un 14% en entrenamiento y en un 9% en test los resultados obtenidos en la aproximación básica WM. Además, elimina cuatro reglas respecto a WM y WM-S, lo que representa un 17% del conjunto de reglas candidatas obtenidas con WM. El modelo que obtiene menor número de reglas es WM-S, pero su funcionamiento es peor que el modelo original obtenido con WM.

Considerando la aproximación DC, el modelo DC-WS obtiene de nuevo el mejor funcionamiento, mejorando en un 37% y un 31% en entrenamiento y prueba respectivamente, los resultados obtenidos por DC. Si nos fijamos en DC-S observamos que obtiene un comportamiento similar a DC-WS en el error de prueba, pero en el error de entrenamiento su comportamiento es significativamente peor. Por otro lado, DC-W tiene un comportamiento similar a DC-WS en el error de entrenamiento, pero el error de prueba es significativamente peor. El modelo DC-WS elimina 6 reglas respecto a DC, lo que supone un 25% del conjunto de reglas candidatas obtenidas con DC. Además, nuestro modelo tiene solo 3 reglas con doble consecuente respecto a las 10 reglas con doble consecuente consideradas en DC y DC-W, lo que simplifica bastante el modelo obtenido.

Según los resultados de la Tabla II podemos decir que la técnica propuesta es más robusta que si consideramos reglas ponderadas y selección de reglas de manera independiente. Además, W y S solo obtienen buenos resultados considerando WM y DC respectivamente.

Las tablas de decisión de los modelos obtenidos por los métodos WM-W, WM-WS, DC-W y DC-WS se presentan en la Figura 3. Cada celda de las tablas representa un subespacio difuso/regla y contiene el consecuente(s) de salida —la primera

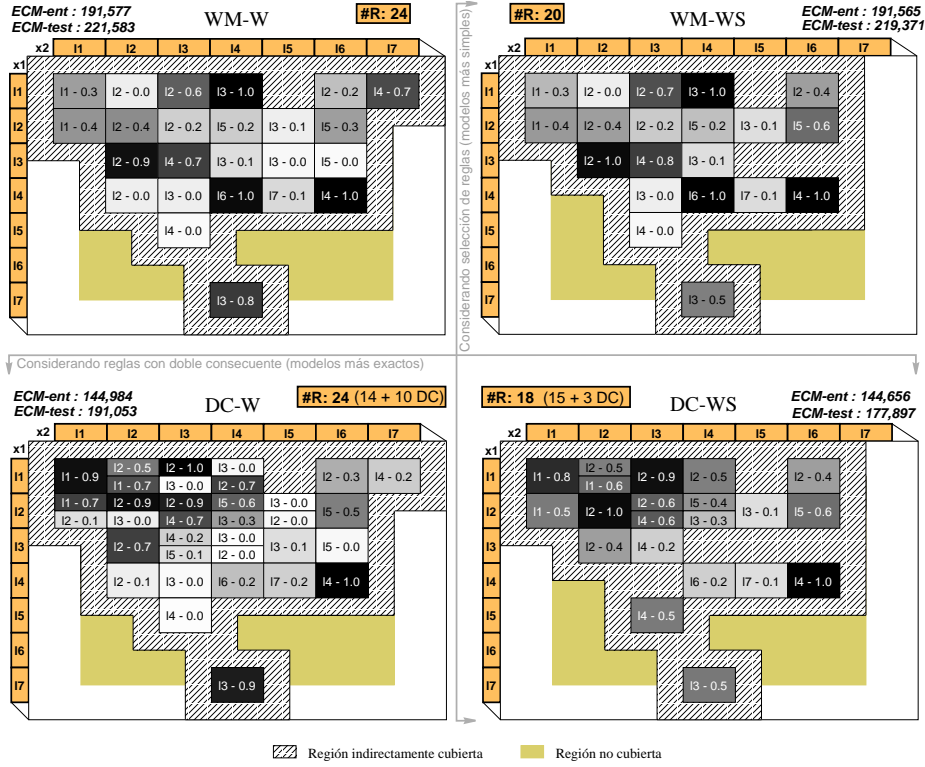


Fig. 3. Tablas de decisión de los modelos obtenidos considerando los enfoques WM y DC.

y/o segunda en importancia cuando consideramos la aproximación DC—, la(s) etiqueta(s) correspondiente(s) y el peso asociado cuando este sea considerado. Los pesos han sido representados gráficamente mediante la escala de grises, donde (1.0) es negro y (0.0) es blanco. De esta manera podemos ver fácilmente la importancia de la regla respecto a sus vecinas y ayuda al experto a identificar las reglas importantes.

Podemos destacar que WM-WS presenta prácticamente los mismos pesos que WM-W. Podemos observar que algunas de las reglas que aparecen en WM-W con un peso próximo a 0 han sido eliminadas en la combinación de pesos y selección de reglas. Sin embargo, no se han eliminado todas, ya que como dijimos en la sección II, éstas son las únicas de su región (l_5-l_3) o mejoran el funcionamiento del sistema interactuando a bajo nivel (l_1-l_2 y l_4-l_3). Si nos centramos ahora en la aproximación DC podemos observar que los pesos vuelven a ser similares en DC-WS y DC-W. Sin embargo, podemos observar algunas diferencias en las tablas debido al gran número de reglas con doble consecuente que hay en DC-W.

Teniendo en cuenta los resultados obtenidos, tanto con WM como con DC, podríamos decir que WS hereda la precisión de las reglas ponderadas y la simplificación de la selección de reglas.

VII. CONCLUSIONES

En este trabajo hemos propuesto el uso de reglas difusas ponderadas combinado con la selección de reglas para obtener modelos difusos lingüísticos más simples y precisos. Para ello hemos utilizado un AG que codifica las reglas y los pesos en cada cromosoma y cuyo objetivo es obtener modelos lingüísticos simples y precisos, pero manteniendo la interpretabilidad a un nivel aceptable.

De acuerdo con los resultados obtenidos, la aproximación propuesta hereda la característica de precisión de las reglas ponderadas y la simplicidad de la selección de reglas. Esto es debido a:

- la habilidad de los pesos de las reglas para indicar el nivel de interacción de cada una de las reglas con el resto, mejorando el funcionamiento global del modelo difuso ponderado.
- la complementariedad de las características que presentan el uso de pesos en las reglas y el proceso de selección de reglas. La habilidad de la selección de reglas para reducir el espacio de búsqueda (eligiendo solo las reglas que cooperen bien) se combina con la mejora en la capacidad de cooperación de las reglas (debido al apropiado nivel de interacción entre las reglas seleccionadas) producida por el uso de pesos en ellas.

REFERENCIAS

- [1] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena, Eds., *Accuracy improvements in linguistic fuzzy modeling*, vol. 129, Springer-Verlag, Heidelberg, Germany, 2003.
- [2] S. Chiu, "Fuzzy model identification based on cluster estimation," *Journal of Intelligent and Fuzzy Systems*, vol. 2, pp. 267–278, 1994.
- [3] S. Halgamuge, M. Glesner, "Neural networks in designing fuzzy systems for real world applications," *Fuzzy Sets and Systems*, vol. 65, no. 1, pp. 1–12, 1994.
- [4] H. Ishibuchi, K. Nozaki, N. Yamamoto, H. Tanaka, "Selecting fuzzy if-then rules for classification problems using genetic algorithms," *IEEE Transactions on Fuzzy Systems*, vol. 9, no. 3, pp. 260–270, 1995.
- [5] A. Krone, H. Krause, T. Slawinski, "A new rule reduction method for finding interpretable and small rule bases in high dimensional search spaces," in *Proceedings of the 9th IEEE International Conference on Fuzzy Systems*, San Antonio, TX, USA, 2000, pp. 693–699.
- [6] R. Rovatti, R. Guerrieri, G. Baccarani, "Fuzzy rules optimization and logic synthesis," in *Proceedings of the 2nd IEEE International Conference on Fuzzy Systems*, San Francisco, USA, 1993, vol. 2, pp. 1247–1252.
- [7] M. Setnes, R. Babuska, U. Kaymak, H. R. van Nauta-Lemke, "Similarity measures in fuzzy rule base simplification," *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, vol. 28, pp. 376–386, 1998.
- [8] M. Setnes, H. Hellendoorn, "Orthogonal transforms for ordering and reduction of fuzzy rules," in *Proceedings of the 9th IEEE International Conference on Fuzzy Systems*, San Antonio, Texas, USA, 2000, vol. 2, pp. 700–705.
- [9] J. Yen, L. Wang, "Simplifying fuzzy rule-based models using orthogonal transformation methods," *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, vol. 29, pp. 13–24, 1999.
- [10] O. Cordón, F. Herrera, "A proposal for improving the accuracy of linguistic modeling," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 4, pp. 335–344, 2000.
- [11] K. Nozaki, H. Ishibuchi, H. Tanaka, "A simple but powerful heuristic method for generating fuzzy rules from numerical data," *Fuzzy Sets and Systems*, vol. 86, pp. 251–270, feb 1997.
- [12] J. S. Cho, D. J. Park, "Novel fuzzy logic control based on weighting of partially inconsistent rules using neural network," *Journal of Intelligent Fuzzy Systems*, vol. 8, pp. 99–110, 2000.
- [13] H. Ishibuchi, T. Takashima, "Effect of rule weights in fuzzy rule-based classification systems," *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 3, pp. 260–270, 2001.
- [14] N. R. Pal, K. Pal, "Handling of inconsistent rules with an extended model of fuzzy reasoning," *Journal of Intelligent Fuzzy Systems*, vol. 7, pp. 55–73, 1999.
- [15] R. Alcalá, J. Casillas, O. Cordón, F. Herrera, "Improvement to the cooperative rules methodology by using the ant colony system algorithm," *Mathware & Soft Computing*, vol. 8, no. 3, pp. 321–335, 2001.
- [16] P. Thrift, "Fuzzy logic synthesis with genetic algorithms," in *Proceedings of the 4th International Conference on Genetic Algorithms (ICGA'91)*, San Mateo, CA, 1991, pp. 509–513, Morgan Kaufmann.
- [17] J. H. Holland, *Adaptation in Natural and Artificial Systems*, Ann arbor: The University of Michigan Press, 1975.
- [18] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag, 1996.
- [19] L. X. Wang, J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 6, pp. 1414–1427, Nov. 1992.
- [20] W. Yu, Z. Bien, "Design of fuzzy logic controller with inconsistent rule base," *Journal of Intelligent Fuzzy Systems*, vol. 2, pp. 147–159, 1994.
- [21] O. Cordón, F. Herrera, A. Peregrín, "Applicability of the fuzzy operators in the design of fuzzy logic controllers," *Fuzzy Sets and Systems*, vol. 86, pp. 15–41, Feb. 1997.
- [22] W. E. Combs, J. E. Andrews, "Combinatorial rule explosion eliminated by a fuzzy rule configuration," *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 1, pp. 1–11, 1998.
- [23] F. Herrera, M. Lozano, J. L. Verdegay, "A learning process for fuzzy control rules using genetic algorithms," *Fuzzy Sets and Systems*, vol. 100, pp. 143–158, 1998.
- [24] O. Cordón, F. Herrera, "A three-stage evolutionary process for learning descriptive and approximative fuzzy logic controller knowledge bases from examples," *International Journal of Approximate Reasoning*, vol. 17, no. 4, pp. 369–407, 1997.
- [25] O. Cordón, M. J. del Jesús, F. Herrera, "Genetic learning of fuzzy rule-based classification systems cooperating with fuzzy reasoning methods," *International Journal of Intelligent Systems*, vol. 13, no. 10–11, pp. 1025–1053, 1998.
- [26] H. Ishibuchi, T. Murata, I. B. Türksen, "Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems," *Fuzzy Sets and Systems*, vol. 89, pp. 135–150, 1997.
- [27] T. C. Chin, X. M. Qi, "Genetic algorithms for learning the rule base of fuzzy logic controller," *Fuzzy Sets and Systems*, vol. 97, no. 1, pp. 1–7, 1998.
- [28] A. F. Gómez-Skarmeta, F. Jiménez, "Fuzzy modeling with hybrid systems," *Fuzzy Sets and Systems*, vol. 104, pp. 199–208, 1999.
- [29] H. Roubos, M. Setnes, "Compact fuzzy models through complexity reduction and evolutionary optimization," *Proceedings of the 9th IEEE International Conference on Fuzzy Systems (San Antonio, Texas, USA)*, vol. 2, pp. 762–767, 2000.
- [30] J. E. Baker, "Reducing bias and inefficiency in the selection algorithm," in *Proceedings of the 2nd International Conference on Genetic Algorithms (ICGA'87)*, J. J. Grefenstette, Ed., Hillsdale, NJ, USA, 1987, pp. 14–21, Lawrence Erlbaum Associates.
- [31] F. Herrera, M. Lozano, J. L. Verdegay, "Tuning fuzzy controllers by genetic algorithms," *International Journal of Approximate Reasoning*, vol. 12, pp. 299–315, 1995.
- [32] O. Cordón, F. Herrera, L. Sánchez, "Solving electrical distribution problems using hybrid evolutionary data analysis techniques," *Applied Intelligence*, vol. 10, pp. 5–24, 1999.