Universidad de Granada

Departamento de Ciencias de la Computación
e Inteligencia Artificial

Optimización Evolutiva Multi-Objetivo de Medidas de Complejidad e Interpretabilidad Semántica para Sistemas Basados en Reglas Lingüísticas

Tesis Doctoral

María José Gacto Colorado

Granada, Septiembre de 2010

Editor: Editorial de la Universidad de Granada
Autor: María José Gacto Colorado
D.L.: GR 4046-2011
ISBN: 978-84-694-6679-7

UNIVERSIDAD DE GRANADA

Optimización Evolutiva Multi-Objetivo de Medidas de Complejidad e Interpretabilidad Semántica para Sistemas Basados en Reglas Lingüísticas

MEMORIA QUE PRESENTA

María José Gacto Colorado

PARA OPTAR AL GRADO DE DOCTOR EN INFORMÁTICA

Septiembre de 2010

DIRECTORES

Rafael Alcalá Fernandez y Francisco Herrera Triguero

Departamento de Ciencias de la Computación
e Inteligencia Artificial

La memoria titulada "*Optimización Evolutiva Multi-Objetivo de Medidas de Complejidad e Interpretabilidad Semántica para Sistemas Basados en Reglas Lingüísticas*", que presenta D. María José Gacto Colorado para optar al grado de doctor, ha sido realizada dentro del programa de doctorado "*Diseño, Análisis y Aplicaciones de Sistemas Inteligentes*" del Departamento de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada bajo la dirección de los doctores D. Rafael Alcalá Fernandez y D. Francisco Herrera Triguero.

Granada, Septiembre de 2010

El Doctorando — Los Directores

Fdo: María José Gacto Colorado — Fdo: Rafael Alcalá Fernandez — Fdo: Francisco Herrera Triguero

Agradecimientos

Dedico esta memoria a mi familia y amigos, en especial a mi marido. Sin su gran apoyo en los momentos más difíciles no habría sido posible la realización de esta tesis. Gracias por darme fuerzas para seguir adelante y por convertirte en alguien imprescindible. Espero que a partir de ahora los dos podamos empezar a disfrutar un poco más de la vida.

En primer lugar quiero agradecer a mis padres, mis suegros, mi hermana, mis cuñados Juan de Dios y Jesús, mi cuñada Zori y mis sobrinos por su cariño y paciencia por no haberles podido dedicar todo el tiempo que me hubiera gustado. Igualmente, doy las gracias a todos mis amigos de Córdoba y de Granada por escucharme y animarme.

Agradezco a mis directores, Rafael Alcalá y Francisco Herrera, el valioso tiempo dedicado y el cariño con el que siempre me han tratado. Sé que están muy ocupados, pero siempre han tenido tiempo para mí. Ambos han hecho posible la existencia de esta tesis doctoral con su inestimable ayuda.

Por último, quiero dar las gracias a mis compañeros de la Universidad de Jaén que con los que he dado mis primeros pasos como profesora en la universidad, especialmente a todos los profesores con los que he compartido asignaturas, gracias a su ayuda he podido ejercer esta profesión tan gratificante y espero poder seguir aprendido de ellos durante mucho tiempo. También quiero expresar mi gratitud al grupo EC3 de Biblioteconomía con los que durante tantos años he trabajado y tantas horas he compartido.

GRACIAS A TODOS

Índice

Parte I. Memoria

1. Introducción

Nuestro interés en esta memoria reside en el estudio de la interpretabilidad de los Sistemas Basados en Reglas Difusas (SBRDs) Lingüísticos para el caso de problemas de regresión, buscando obtener no sólo un buen equilibrio entre dos objetivos contradictorios como son precisión e interpretabilidad, sino también determinar qué medidas pueden utilizarse para cuantificar la interpretabilidad de los SBRDs lingüísticos. Para ello proponemos usar Algoritmos Evolutivos Multi-Objetivo (AEMOs), que permiten generar frentes de Pareto con distintos equilibrios para ambos objetivos. Esta forma de trabajar permite no sólo seleccionar la solución que más nos interese en cada momento, sino centrar la búsqueda en la zona del frente más prometedora.

Para llevar a cabo este estudio, la presente memoria se divide en dos partes, la primera de ellas dedicada al planteamiento del problema y discusión de los resultados y la segunda correspondiente a las publicaciones asociadas al estudio.

En la Parte I de la memoria comenzamos con una sección dedicada al "Planteamiento" del problema, introduciendo éste con detalle y describiendo las técnicas utilizadas para resolverlo. Asimismo, definimos los problemas abiertos en este marco de trabajo que justifican la realización de esta memoria así como los objetivos propuestos. Posteriormente, incluimos una sección de "Discusión de Resultados", que proporciona una información resumida de las propuestas y los resultados más interesantes obtenidos en las distintas partes en las que se divide el estudio. La sección "Comentarios Finales" resume los resultados obtenidos en esta memoria y presenta algunas conclusiones sobre éstos, para finalmente comentar algunos aspectos sobre trabajos futuros que quedan abiertos en la presente memoria.

Por último, para desarrollar los objetivos planteados, la Parte II de la memoria está constituida por cinco publicaciones distribuidas en cuatro partes:

- Mejora de Controladores Difusos Obtenidos a partir de Expertos: Un Caso de Estudio sobre un Sistema de Ventilación, Calefacción y Aire Acondicionado - *Improving Fuzzy Logic Controllers Obtained by Experts: A Case Study in HVAC Systems.*

- Interpretabilidad de los Sistemas Basados en Reglas Difusas Lingüísticos: Una Revisión sobre Medidas de Interpretabilidad - *Interpretability of Linguistic Fuzzy Rule-Based Systems: An Overview on Interpretability Measures*

- Algoritmos Evolutivos Multi-Objetivo que Combinan las Técnicas de Ajuste y de Selección de Reglas para Obtener Sistemas Basados en Reglas Difusas Lingüísticos Precisos y Compactos (con dos de las cinco publicaciones) - *Multi-objective Genetic Algorithms for Tuning and Rule Selection to Obtain Accurate and Compact Linguistic Fuzzy Rule-Based Systems*
- Integración de un Índice para Preservar la Interpretabilidad Semántica en la Selección y Ajuste Evolutivos Multi-Objetivo de los Sistemas Difusos Lingüísticos - *Integration of an Index to Preserve the Semantic Interpretability in the Multi-Objective Evolutionary Rule Selection and Tuning of Linguistic Fuzzy Systems*

1.1. Planteamiento

El Modelado de Sistemas es una de las aplicaciones más importantes en el campo de los SBRDs [Mam74, MA75, PdO96, Zad73]. El Modelado Difuso puede considerarse como un enfoque utilizado para modelar un sistema haciendo uso de un lenguaje descriptivo basado en la Lógica Difusa [Zad65, Zad73] con predicados difusos [SY93]. Este tipo de modelado esta caracterizado principalmente por dos propiedades, que permiten asegurar la calidad del modelo difuso obtenido:

- *Precisión*: Es la capacidad de representar fielmente un sistema real. Un sistema será mejor cuanto mayor similaridad exista entre la respuesta del sistema real y el modelo difuso.
- *Interpretabilidad*: Es la capacidad de expresar el comportamiento del sistema de una manera entendible. Ésta es una propiedad subjetiva que depende de varios factores, principalmente la estructura del modelo, el número de variables de entrada, el número de reglas difusas, el número de términos lingüísticos, la forma de los conjuntos difusos, etc.. No existe una medida estándar para evaluar como de buena es la interpretabilidad de un sistema.

Dependiendo del objetivo principal que se desee satisfacer, podemos distinguir entre dos clases de Modelado Difuso:

- *Modelado Difuso Lingüístico (MDL)* — Este tipo de modelado se realiza generalmente por medio de SBRD lingüísticos, también conocidos como de tipo Mamdani [Mam74, MA75]. En este caso, el principal requisito es la interpretabilidad, y el concepto de variable lingüística [Zad75] desempeña un papel fundamental. Además de su gran facilidad para interpretar el comportamiento del sistema, su estructura proporciona un marco natural para incluir conocimiento experto, por lo que son los más empleados en la actualidad.

  Los SBRDs lingüísticos están formados por reglas con la siguiente estructura:

  $$\textbf{SI } X_1 \text{ es } A_1 \text{ y } \ldots \text{ y } X_n \text{ es } A_n \textbf{ ENTONCES } Y \text{ es } B,$$

  donde X_i (Y) son las variables lingüísticas de entrada (salida), y A_i y B las etiquetas lingüísticas con los conjuntos difusos μ_{A_i} y μ_B asociados definiendo su significado. Estos términos lingüísticos se toman de una *semántica global* que define la gama posible de conjuntos difusos usados en cada variable.
- *Modelado Difuso Preciso (MDP)* — En esta clase de modelado se persigue principalmente la precisión de los modelos obtenidos, dejando a un lado su legibilidad. Para ello, podemos emplear los SBRDs aproximativos, que se caracterizan por el uso directo de variables difusas.

Así, cada regla difusa presenta su propia semántica, es decir, las variables toman diferentes conjuntos difusos como valores en lugar de términos lingüísticos. Dado que en los SBRDs aproximativos no se emplea una semántica global, los conjuntos difusos no pueden interpretarse con facilidad. Estos modelos pretenden ser más precisos que los anteriores, es decir, capturan la información del problema de un modo más exacto a costa de la consiguiente pérdida de interpretabilidad.

La estructura de regla considerada es la siguiente:

$$\textbf{SI } X_1 \text{ es } \widehat{A}_1 \text{ y } \ldots \text{ y } X_n \text{ es } \widehat{A}_n \textbf{ ENTONCES } Y \text{ es } \widehat{B},$$

donde $\widehat{A}_i$ y $\widehat{B}$ son conjuntos difusos sin una interpretación lingüística directa.

En esta tesis, nos hemos centrado en los SBRD lingüísticos que por su propia naturaleza son más interpretables que los aproximativos.

1.1.1. Sistemas Basados en Reglas Difusas Lingüísticos

Los SBRDs lingüísticos fueron inicialmente propuestos por Mamdani y Assilian [Mam74, MA75], que plasmaron las ideas preliminares de Zadeh [Zad73] en el primer SBRD diseñado para una aplicación de control. Este tipo de sistemas difusos es uno de los más usados desde entonces y se conoce también por el nombre de SBRD de tipo Mamdani o, sencillamente, controlador difuso, ya que su aplicación principal ha sido históricamente el control de sistemas. La Figura 1 muestra la estructura de un SBRD lingüístico.

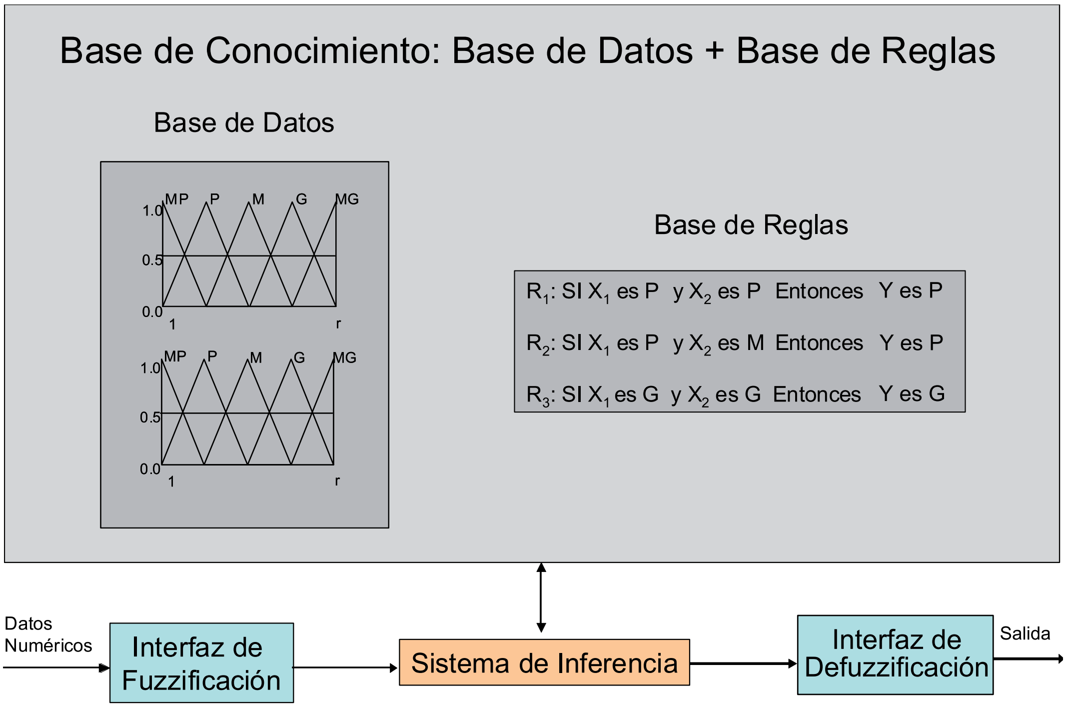


Figura 1: Estructura de un SBRD lingüístico

Los SBRDs lingüísticos están compuestos por dos componentes fundamentales que son la *Base de Conocimiento* (BC) y el módulo con el motor de inferencia. La BC de un SBRDs lingüístico se puede dividir en dos partes, la *Base de Datos* (BD) y la *Base de Reglas* (BR):

- La BD, contiene los términos lingüísticos considerados en las reglas lingüísticas y las funciones de pertenencia que definen la semántica de las etiquetas difusas. De este modo, cada variable lingüística incluida en el problema tendrá asociada una partición difusa asociada con cada uno de sus términos lingüísticos. La Figura 2 muestra un ejemplo de una partición difusa con cinco etiquetas.

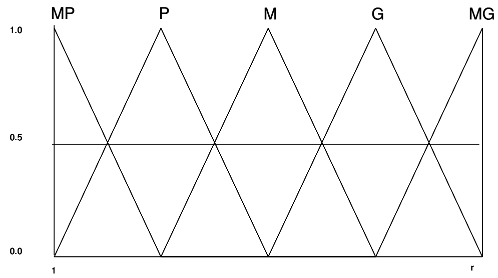


Figura 2: Ejemplo de una partición difusa

  Ésto puede ser considerado como una aproximación a la discretización para dominios continuos donde establecemos un grado de pertenencia a los items (etiquetas), donde hemos de incluir un solapamiento entre ellos, y el motor de inferencia maneja el emparejamiento entre los patrones y las reglas proporcionando una salida acorde a los consecuentes de las reglas con un emparejamiento positivo. La determinación de las particiones difusas es crucial en modelado difuso, especialmente en los problemas de control y de regresión, y la granularidad de las particiones difusas juega un papel importante para el comportamiento del SBRDs.

- La BR, que está compuesta por una colección de reglas lingüísticas que se unen mediante una conectiva de reglas (operador "también"). En el caso de un SBRDs, se pueden disparar múltiples reglas simultáneamente con la misma entrada.

El módulo con el motor de inferencia incluye:

- Un *interfaz de fuzzificación*, que se encarga de transformar los datos de entrada precisos en valores utilizables en el proceso de razonamiento difuso, es decir, en algún tipo de conjunto difuso.
- Un *sistema de inferencia*, que a través de los datos recibidos por el interfaz de fuzzificación, utiliza la información contenida en la BC para llevar a cabo el proceso de inferencia difuso.
- Un *interfaz de defuzzificación*, que transforma la acción difusa resultante del proceso de inferencia en una acción precisa que constituye la salida global del SBRDs.

1.1.2. Interpretabilidad de los Sistemas Basados en Reglas Difusas Lingüísticos

En esta sección se examinan algunas ideas básicas y trabajos en el tema de la interpretabilidad del modelado difuso lingüístico. Junto con la revisión en [MF08], que representan la mayoría de los trabajos existentes en la literatura especializada, un marco para clasificar la interpretabilidad del modelado difuso en alto nivel de interpretabilidad y bajo nivel interpretabilidad ha sugerido recientemente en [ZG08]:

- Alto nivel de Interpretabilidad se obtiene en el nivel de reglas difusas mediante la reducción de la complejidad global en términos de algunos criterios, tales como un número moderado de variables y reglas, completitud y consistencia de las reglas (complexity-based interpretability).
- Bajo nivel de Interpretabilidad se alcanza a nivel de conjuntos difusos mediante la optimización de las funciones de pertenencia en términos de los criterios semánticos sobre las funciones de pertenencia (semantics-based interpretability).

Las técnicas de reducción de la complejidad que se utilizan en el modelado tradicional de sistemas pueden servir para la optimización de reglas difusas, que se corresponde con el objetivo de la parsimonia de la BR difusa, uno de los principales criterios de interpretabilidad de alto nivel de los sistemas difusos. Esta aclaración es útil, ya que hay abundantes métodos tradicionales de modelado de sistemas para la reducción de la complejidad con un gran potencial de ser utilizadas para inducir una BR compacta en la modelización de sistemas difusos. En los primeros trabajos de Ishibuchi y otros [INYT95, IMT97] utilizan selección de reglas sobre un conjunto inicial de reglas de clasificación y dos criterios diferentes, la precisión y el número de reglas. Junto con los trabajos [IMT97], [INM01, IY04, IN07] en los que se buscar optimizar los criterios de complejidad para el caso de AEMOs. Además, la longitud de la regla (a veces se utiliza en combinación con el número de reglas) se ha incluido para reducir al mínimo la longitud de las reglas mediante la selección de reglas [CDLM07, INM01, IY04] o el aprendizaje de reglas [INM01, IN07, PK08]. Un método ha sido propuesto también en [AGHAF07] y profundamente discutido en [GAH09] para reducir al mínimo el número de reglas, junto con un ajuste de las funciones de pertenencia.

Bajo nivel de interpretabilidad se logra mediante la optimización de funciones de pertenencia en el nivel de los conjuntos difusos. En concreto, este nivel de interpretabilidad proviene de la mejora de la interpretabilidad mediante la introducción de restricciones o medidas en los criterios semánticos dentro del modelado difuso, que se centran en los cambios de las funciones de pertenencia [ZG08]. Las aproximaciones clásicas como [dO99a, dO99b] definen algunos criterios semánticos útiles como distinguibilidad, moderado número de funciones de pertenencia, posición natural del cero, la normalidad y la cobertura. Estas propiedades fueron incluidas más tarde en un AEMO en [FdOP07] para controlar su interacción cuando evolucionan de forma conjunta. Otros trabajos se han centrado en la definición de métricas de similitud adecuadas, como una forma de medir la distinguibilidad y la cobertura de las funciones de pertenencia [MCF07], que a veces se utilizan para fijar unos valores mínimos de la cobertura [CL00, JvSS99], y otras para definir los valores máximos de similitud para la unión de los conjuntos difusos y reglas [EV00, SBKvNL98] (sobre todo cuando funciones de pertenencia provienen de técnicas de clustering). En [Nau03], una medida de similitud se ha optimizado para buscar una buena cobertura de las funciones de pertenencia junto con dos criterios de complejidad en un índice combinado. Un AEMO es utilizado en [BLMS09] para llevar a cabo la adaptación de contextos. Este algoritmo considera el error del sistema y un índice de interpretabilidad para preservar el orden difuso y una buena distinguibilidad.

Además, algunos otros trabajos tratan de ir un paso por delante al considerar todo este tipo de medidas en un marco lingüístico a fin de buscar una definición más global de interpretabilidad [AMG08, BB03]. En este sentido, Alonso y otros en [AMR09] presenta un marco conceptual para la caracterización de la interpretabilidad de los SBRDs. Hace referencia a los trabajos de [MF08] y [ZG08] que se combinan en varios niveles de interpretabilidad (extendiendo la categorización de bajo-alto nivel).

La mayoría de los enfoques basados en semántica se centran principalmente en la búsqueda de particiones con un buen solapamiento entre las funciones de pertenencia (cobertura y distinguibilidad) usando para ello propiedades absolutas. Debido a que la interpretabilidad está fuertemente relacionada con el contexto del problema y la percepción de los usuarios, las medidas interpretabilidad semántica absolutas no están todavía completamente aceptadas como la mejor manera de preservar la interpretabilidad semántica de los SBRDs lingüísticos.

1.1.3. Sistemas Difusos Evolutivos

Un Sistema Difuso Evolutivo (SDE) [Her08], llamado en inglés Genetic Fuzzy System, es básicamente un sistema difuso mejorado por un proceso de aprendizaje basado en computación evolutiva, que incluye Algoritmos Genéticos (AGs), programación genética y estrategias de evolución, entre otros algoritmos evolutivos.

El aspecto central para el uso de un AG para el aprendizaje automático de un SBRD es que el proceso de diseño de la BC que puede ser analizado como un problema de optimización. Desde el punto de vista de la optimización, encontrar una BC apropiada es equivalente a codificarla como una estructura de parámetros y entonces encontrar los valores de los parámetros que den el óptimo para una función de fitness. Los parámetros de la BC proporcionan el espacio de búsqueda que se trasforma de acuerdo a una representación genética. De este modo, el primer paso en el diseño de un SDE es decidir qué partes de la BC estarán sujetas a la optimización por parte del AG. La Figura 3 muestra la estructura de un SDE.

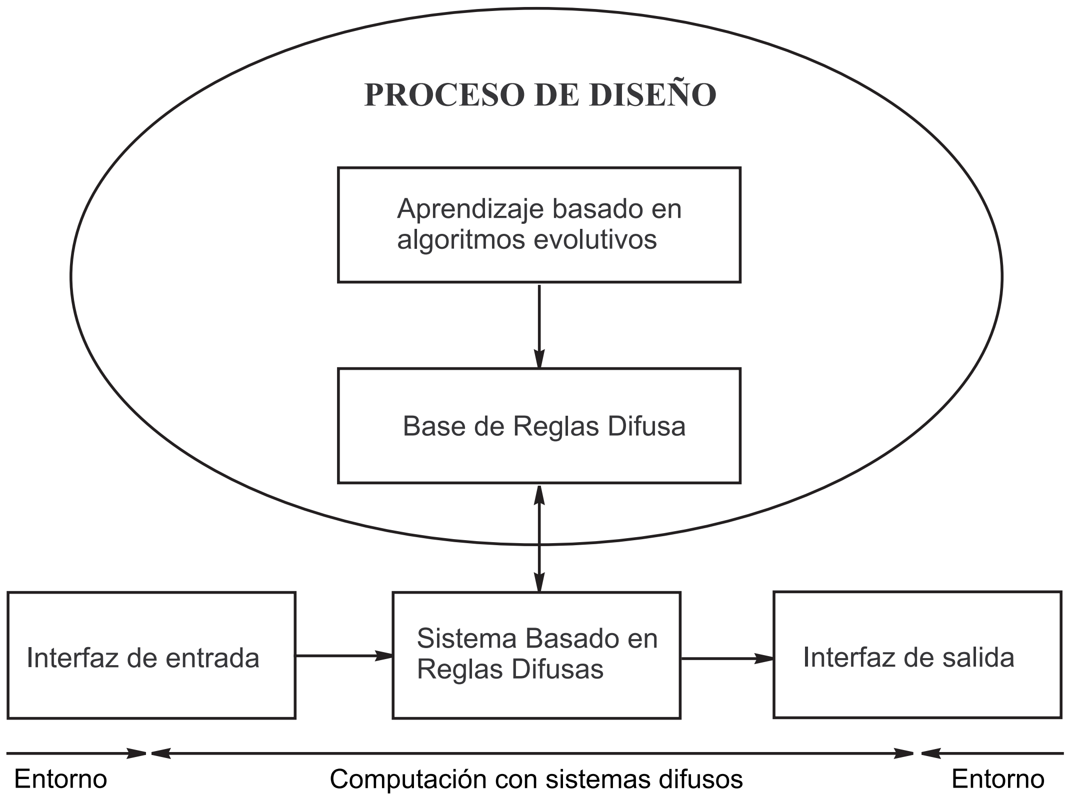


Figura 3: Estructura de un SDE

En los últimos años podemos observar un incremento de los artículos publicados en la materia, debido al alto potencial de los SDEs. Contrariamente a las redes neuronales, clustering, inducción de reglas y muchas otras propuestas de aprendizaje automático, los AGs proporcionan un medio para codificar y evolucionar operadores de agregación en el antecedente de las reglas, diferentes semánticas de las reglas u operadores de agregación de la BR, entre otros. De este modo, los AGs continúan siendo hoy uno de los pocos esquemas de adquisición de conocimiento disponibles para diseñar y, de algún modo, optimizar los SBRDs con respecto a las decisiones de diseño, permitiendo a los usuarios seleccionar qué componentes quedan fijas y cuáles se evolucionan de acuerdo a las medidas de rendimiento.

En primera instancia, las propuestas de SDEs se pueden dividir en dos tipos de procesos: aprendizaje y post-procesamiento. Es difícil realizar una clara distinción entre ambos procesos, dado que establecer una frontera precisa es tan complicado como definir el concepto de aprendizaje. El primer hecho que debemos de tomar en consideración es la existencia o no de una BC previa, incluyendo la BD y la BR. En el entorno de trabajo de los SDEs, se pueden distinguir los siguientes tipos de propuestas siguiendo la taxonomía presentada en [Her08]:

- Aprendizaje genético. La primera posibilidad es aprender los componentes de la BC (donde podemos incluso incluir un motor de inferencia adaptativo). A continuación, describimos las cuatro propuestas que pueden encontrarse dentro del aprendizaje genético:

    1. *Aprendizaje Genético de Reglas.* La mayoría de las aproximaciones propuestas para aprender de forma automática la BC a partir de información numérica se han centrado en el aprendizaje de la BR, utilizando una BD predefinida. El modo usual para definir esta BD requiere escoger un número de términos lingüísticos para cada variable lingüística (un número impar entre 3 y 9, que será normalmente el mismo para todas las variables) y darle el valor a los parámetros del sistema mediante una distribución uniforme de los términos lingüísticos en el universo de discurso de las variables. La propuesta pionera para este tipo de ajuste puede encontrarse en [Thr91].
    2. *Selección Genética de Reglas (como parte del proceso de aprendizaje).* A veces tenemos un gran número de reglas extraídas a través de un método de Minería de Datos que sólo tiene como objetivo la precisión final del modelo sin importar su complejidad. Una BR con un excesivo número de reglas hace difícil comprender el comportamiento del SBRD. Así, podemos encontrar diferentes tipos de reglas en un mismo conjunto de reglas difusas: reglas irrelevantes, reglas redundantes, reglas erróneas y reglas en conflicto, que perturban el rendimiento del SBRD cuando coexisten con otras. Para enfrentarse a este problema se puede utilizar un proceso genético de selección de reglas que obtiene un subconjunto de reglas optimizado a partir de un conjunto de reglas difusas previo. En [INYT95] podemos encontrar la primera contribución en este área.
    3. *Aprendizaje Genético de la BD.* Existe otro modo de generar toda la BC que considera dos procesos diferentes para obtener ambos componentes, es decir, la BD y la BR. El proceso de generación de la BD nos permite aprender la forma de las funciones de pertenencia y otras componentes de la BD como las funciones de escalado o la granularidad de las particiones difusas, entre otros. Este proceso de generación de la BD puede utilizar una medida para evaluar la calidad de la BD, lo que se denominaría “aprendizaje genético a priori de la BD”. La segunda posibilidad es considerar un proceso de aprendizaje genético incrustado donde el proceso de generación de la BD se realiza conjuntamente con el aprendizaje de la BR del siguiente modo: cada vez que se obtiene una BD mediante el proceso de definición de la BD, el método de generación de la BR se usa para obtener las reglas, y se utiliza por tanto algún tipo de medida de error para validar la BC completa que se obtiene. Debemos indicar que este modo de operación requiere un particionamiento del problema de aprendizaje de la BC. En [CHV01], podemos encontrar una propuesta referente al aprendizaje genético incrustado de la BD.

4. *Aprendizaje genético simultáneo de las componentes de la BC.* Otras propuestas pretenden aprender las dos componentes de la BC simultáneamente. Trabajando de este modo, se cuenta con la posibilidad de obtener una BC de mayor calidad, si bien la desventaja en este caso es el incremento del espacio de búsqueda, que hace que el proceso de aprendizaje se vuelva más difícil y lento. En [HM95], podemos encontrar un trabajo que es una referencia de este tipo de proceso de aprendizaje.

- Técnicas de post-procesamiento. Si existe una BC, consiste en aplicar un algoritmo genético para mejorar el rendimiento del SBRD. A continuación enumeramos tres de las posibilidades que pueden ser consideradas siguiendo este modelo:

1. *Ajuste Genético de los parámetros de la BC.* Para poder llevar a cabo esta tarea, se usa un proceso de ajuste a posteriori considerando toda la BC obtenida (la BD preliminar y la BR derivada) para ajustar los parámetros de la función de pertenencia, sin modificar la BR existente. El proceso de ajuste solo modifica la forma de las funciones de pertenencia y no el número de términos lingüísticos en cada partición difusa, que permanece fijo desde el principio del proceso de diseño. El ajuste de los parámetros de las funciones de pertenencia es la forma más común de derivar las funciones de pertenencia y consiste en alterar los valores de los distintos parámetros que las definen para realizar desplazamientos y/o ensanchamientos de los conjuntos difusos. Esto se puede conseguir ajustando directamente cada uno de los parámetros, usando diferentes factores de escala lineales, aplicando desplazamientos laterales y/o de amplitud. Por ejemplo, si consideramos la función de pertenencia triangular de la Figura 4,

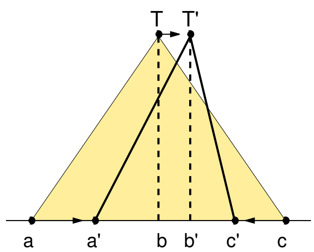


Figura 4: Ajuste clásico de las funciones de pertenencia

alterar los parámetros a, b y c supone variar la forma del conjunto difuso asociado a la función de pertenencia (véase dicha Figura), afectando al comportamiento del SBRD. Lo mismo ocurre en el caso de los demás tipos de funciones de pertenencia (trapezoidales, gaussianas, sigmoidales, etc.). Las técnicas de ajuste de las funciones de pertenencia permiten mejorar el comportamiento del sistema, consiguiendo un aumento de la precisión significativo, sobre todo en problemas de control y regresión.

2. *Sistemas de Inferencia Adaptativos Genéticos.* El principal objetivo de esta propuesta es el uso de expresiones paramétricas en el Sistema de Inferencia, lo que a menudo se denomina Sistemas de Inferencia Adaptativos, para obtener una mayor cooperación entre las reglas difusas y de esta forma modelos difusos más precisos sin perder la interpretabilidad inherente a las reglas lingüísticas. En [AFHMP07, CBFO06, CBM07], podemos encontrar propuestas en este área centradas en regresión y clasificación.

3. *Selección de Reglas (como método de post-procesamiento).* La selección de reglas no sólo se utiliza como parte de un proceso de aprendizaje para obtener un SBRD a partir de datos, sino que también ha sido utilizada como técnica de post-procesamiento para

eliminar reglas redundates o erróneas que pueden aparecer en un sistema previamente obtenido o como consecuencia de un ajuste del modelo. El principal objetivo de esta propuesta es seleccionar un subconjunto de reglas con el mejor nivel de cooperación posible tomando como punto de partida el conjunto inicial, y eliminando aquellas reglas que no afecten o afecten de manera negativa al comportamiento del sistema.

La combinación de técnicas de ajuste de las funciones de pertenencia con métodos de selección de reglas presenta una sinergia positiva cuando dichas técnicas se combinan en el mismo proceso.

1.1.4. Algoritmos Evolutivos Multi-Objetivo

Muchos problemas reales se caracterizan por la existencia de múltiples medidas que deberían optimizarse, o al menos ser satisfechas simultáneamente. Por ejemplo en el diseño de un sistema de control de aire acondicionado, se trata de optimizar un conjunto de parámetros del sistema de control: minimizar el consumo de energía, maximizar el confort de los usuarios, etc...

Por definición, un problema multi-objetivo consiste en Maximizar o Minimizar z donde:

$$z = \ f((x) = (f_1(x), f_2(x), \ldots, f_n(x)).$$

La mejor forma de resolver este tipo de problemas es mediante el uso de los criterios de dominancia y pareto-optimalidad. Las soluciones pareto-optimales o no-dominadas se definen como un vector a que domina a otro b (se nota como $a \preceq b$) si, y solo si (suponiendo maximización):

$$\forall i \in 1, 2, \ldots, n \mid f_i(a) \geq f_i(b) \wedge \exists j \in 1, 2, \ldots, n \mid f_i(a) > f_i(b)$$

Una solución domina a otra si es mejor o igual en todos los objetivos y al menos mejor en uno de ellos. De esta forma, todos los vectores que no son dominados por ningún otro vector se llaman *pareto-optimales o no-dominados*. No suele existir una única solución optimal, existe un conjunto (a veces infinito) de soluciones no-dominadas que forma la frontera o frente del Pareto. En la Figura 5 se muestra un ejemplo de un frente de Pareto para el caso de dos objetivos (Max Q(x), Max T(x)):

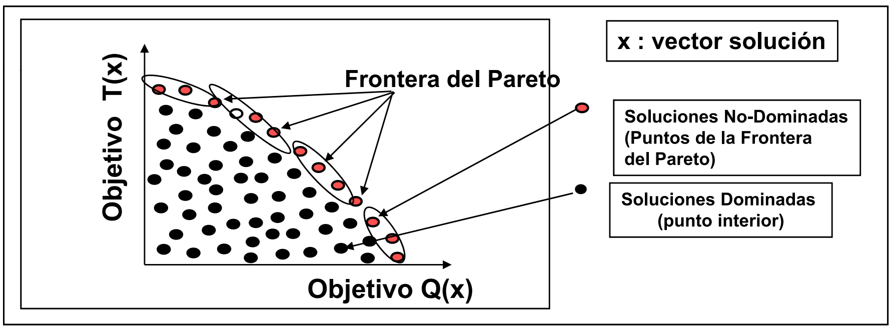


Figura 5: Ejemplo de un frente de Pareto

Los Algoritmos Evolutivos pueden manejar simultáneamente un conjunto de posibles soluciones (población) y permiten encontrar varias soluciones del conjunto de Pareto óptimo en una única

Tabla I.1: Clasificación de los AEMOs

Referencia	**AEMO**	1^a **Gen.**	2^a **Gen.**
[FF93]	MOGA	$\surd$	
[HNG94]	NPGA	$\surd$	
[SD94]	NSGA	$\surd$	
[CT01]	micro-GA		$\surd$
[EMH01]	NPGA 2		$\surd$
[DPAM02]	NSGA-II		$\surd$
[KC00]	PAES		$\surd$
[CKO00, CJKO01]	PESA & PESA-II		$\surd$
[ZT99, ZLT01]	SPEA & SPEA2		$\surd$

ejecución del algoritmo. Además, no son demasiado sensibles a la forma o la continuidad del frente de Pareto (por ejemplo, pueden fácilmente tratar con frentes de Pareto discontinuos y cóncavos).

El primer indicio sobre la posibilidad de utilizar algoritmos evolutivos para resolver un problema multi-objetivo aparece en una tesis de 1967 [Ros67] en la que, sin embargo, no se propuso un AEMO real (el problema multi-objetivo se redefine como un problema de un sólo objetivo y es resuelto con un algoritmo genético mono-objetivo). David Schaffer es considerado como el primero en haber diseñado un AEMO a mediados de la década de los 80 [Sch85]. Schaffer presenta un algoritmo llamado "Vector Evaluated Genetic Algorithm" (VEGA), que consiste en un algoritmo genético simple con un mecanismo de selección modificado. Sin embargo, VEGA tiene una serie de problemas siendo el principal de ellos su incapacidad para mantener soluciones con un rendimiento aceptable, tal vez por encima del promedio, pero no excepcional para cualquiera de las funciones objetivo.

Después de VEGA, los investigadores diseñaron una primera generación de AEMOs caracterizados por su sencillez, donde la principal característica es que combinan un buen mecanismo para seleccionar los individuos no dominados (en algunos casos, basado en el concepto de Pareto optimalidad) con un buen mecanismo para mantener la diversidad. Los AEMOs más representativos de esta generación son los siguientes: "*Nondominated Sorting Genetic Algorithm*" (NSGA) [SD94], "*Niched-Pareto Genetic Algorithm*" (NPGA) [HNG94] y "*Multi-Objective Genetic Algorithm*" (MOGA) [FF93].

La segunda generación de AEMOs comenzó cuando el elitismo se convirtió en un mecanismo estándar tras ser propuesto en SPEA [ZT99]. De hecho, el uso del elitismo es un requisito teórico para garantizar la convergencia de un AEMO. Muchos de los AEMOs que se propusieron en esta segunda generación todavía sobreviven hoy en día. La mayoría de los investigadores coinciden en que algunos de estos enfoques se han adoptado como algoritmos de referencia. De esta manera, "Strength Pareto Evolutionary Algorithm 2" (SPEA2) [ZLT01] y "Nondominated Sorting Genetic Algorithm II" (NSGA-II) [DPAM02] puede ser considerados como los AEMOs más representativos de la segunda generación. En la Tabla I.1 se muestra un resumen de los AEMOs más representativas de ambas generaciones.

Los AEMOs permiten obtener un frente de Pareto con una variedad de soluciones para cada uno de los objetivos considerados. Ésto permite elegir una solución de compromiso entre los distintos objetivos, algo que es muy útil para el caso de objetivos contradictorios, como en el caso del problema del equilibrio entre interpretabilidad y precisión.

Por último, hay que señalar que hoy en día NSGA-II es el paradigma de AEMO para la comunidad científica debido al potencial del operador de "crowding" que este algoritmo utiliza y que por lo general permite obtener un Pareto muy amplio en una gran variedad de problemas, que es una propiedad muy apreciada en este marco de trabajo. De esta manera, la pregunta es: ¿Es NSGA-II el mejor AEMO para obtener el deseado equilibrio entre precisión e interpretabilidad de los SBRDs?. En esta memoria se estudian los dos algoritmos más representativos, mostrando que en el caso de nuestro problema es necesario desarrollar algoritmos específicos a partir de las versiones estándar para manejar el complejo espacio de búsqueda de manera adecuada.

1.2. Justificación

Una vez que se han introducido los principales conceptos a los que se refiere esta memoria, nos planteamos una serie de problemas abiertos que nos sitúan en el planteamiento y la justificación del presente proyecto de tesis.

- Uno de los mayores problemas que existen dentro del campo de la interpretabilidad es que no existe una medida global aceptada que combine medidas de complejidad con medidas de interpretabilidad semántica para medir la interpretabilidad en los SBRD lingúisticos. Al contrario, de lo que ocurre con las medidas de precisión que son bien conocidas y aceptadas.
- Igualmente, no existen medidas de interpretabilidad semántica ampliamente aceptadas en el área de los SBRDs lingüísticos, por lo que habría que definir nuevas medidas que permitan cuantificar la interpretabilidad semántica respecto a unas particiones iniciales (que pueden proceder de un experto).
- En la literatura para el caso específico de problemas de regresión, no existen AEMOs que combinen medidas de precisión junto a medidas de interpretabilidad (medidas de complejidad y/o medidas de interpretabilidad semántica). Por este motivo sería interesante que se definan nuevos algoritmos que puedan manejar esas medidas de forma efectiva.

Como resultado, existen pocos trabajos en la literatura especializada que tratan el tema de la interpretabilidad en los SBRDs lingüísticos para problemas de regresión, en los que es de vital importancia aplicar técnicas de ajuste o aprendizaje de las funciones de pertenencia. Este aspecto complica la obtención de modelos óptimos, no solo precisos sino también altamente interpretables.

1.3. Objetivos

Para dar solución a los distintos problemas que se acaban de mencionar en la sección anterior, la presente memoria se organiza en torno a cuatro grandes objetivos que involucran el estudio del comportamiento de los AEMOs considerando medidas de precisión, complejidad e interpretabilidad semántica de los SBRD lingüísticos en problemas de regresión. En concreto, los objetivos que proponemos son:

- Estudiar la combinación de las técnicas de post-procesamiento de ajuste de las funciones de pertenencia con métodos de selección de reglas en problemas complejos. Combinar ambas técnicas con el objetivo de estudiar su comportamiento como base de la propuesta de algoritmos más avanzados que permitan decrementar la complejidad sin perder la precisión.

- Realizar un estudio sobre las medidas de interpretabilidad existentes en la literatura en el campo de los SBRDs lingüísticos. Proponer una taxonomía que permita analizar qué medidas serían las más interesantes según la parte de la BC a optimizar, en nuestro caso las funciones de pertenencia y las reglas a eliminar.
- Realizar un estudio desde el punto de vista de la complejidad sobre los modelos de ajuste y selección de reglas, desarrollando distintos algoritmos con el objetivo de mejorar el balance entre precisión y simplicidad mediante el ajuste de las funciones de pertenencia y la reducción del número de reglas, respectivamente. Para ello, propondremos el uso de algoritmos multi-objetivo considerando dos objetivos contradictorios: minimizar el error del modelo (precisión) y minimizar el número de reglas (complejidad). Desarrollar los algoritmos adecuados para encontrar los grados de equilibrio entre complejidad y precisión, minimizando dos objetivos con distinto nivel de dificultad que pueden llevar a resultados no deseados (tendencia hacia el objetivo más facil).
- Proponer una medida de calidad desde el punto de vista de la interpretabilidad semántica de las funciones de pertenencia, que junto con el criterio correspondiente de precisión e incluso junto a las medidas para reducir la complejidad, se puedan optimizar mediante un algoritmo multi-objetivo, dando lugar a modelos con una información más útil. Desarrollar distintos algoritmos multi-objetivo que consideren ambos tipos de medidas en los métodos de selección y ajuste de funciones de pertenencia, estudiando los grados de equilibrio entre las distintas medidas.

2. Discusión de Resultados

Esta sección muestra un resumen de las distintas propuestas que se recogen en la presente memoria y presenta una breve discusión sobre los resultados obtenidos por cada una de ellas.

2.1. Mejora de Controladores Difusos Obtenidos a partir de Expertos: Un Caso de Estudio sobre un Sistema de Ventilación, Calefacción y Aire Acondicionado

En este trabajo se proponen varios algoritmos genéticos para resolver un problema complejo de control de un sistema de Ventilación, Calefacción y Aire Acondicionado –en inglés *Heating, Ventilating, and Air Conditioning* (HVAC)–. Para ello se usan técnicas de ajuste (lateral [AAFH07] y lateral con amplitud [AAFGH07]) combinadas con selección de reglas para mejorar el comportamiento de controladores logicos difusos obtenidos a partir de expertos en problemas complejos. Este trabajo muestra un primer estudio de la combinación de las técnicas de ajuste y selección. El artículo asociado a esta parte es:

- R. Alcalá, J. Alcalá-Fdez, M.J. Gacto, F. Herrera, Improving Fuzzy Logic Controllers Obtained by Experts: A Case Study in HVAC Systems. Applied Intelligence 31:1 (2009) 15-30, doi:10.1007/s10489-007-0107-6. Citado en 2 ocasiones.

Este trabajo es el punto de partida que justifica la necesidad de técnicas que traten de decrementar la complejidad sin perder precisión, porque en los resultados se vé que en algunas ocasiones se encuentran soluciones con menos reglas y al mismo tiempo con mejor precisión. Dichos resultados muestran que el espacio de búsqueda es altamente complejo, puesto que las soluciones con mayor precisión no tienen por qué ser las más complejas. Sin embargo, parece necesario el desarrollo de algoritmos específicos que se centren en la búsqueda de dichas soluciones.

2.2. Interpretabilidad de los Sistemas Basados en Reglas Difusas Lingüísticos: Una Revisión sobre Medidas de Interpretabilidad

En este trabajo se presenta una revisión de las distintas medidas de interpretabilidad existentes en la literatura y se propone una taxonomía como forma de organizar las medidas en cuatro cuadrantes. La taxonomía propuesta esta basada en un doble eje:

- Complejidad frente a Interpretabilidad Semántica.
- Base de Reglas frente a Particionamiento Difuso.

Esta taxonomía propuesta proviene de la combinación de ambos ejes y conduce a la aparición de los cuadrantes siguientes dedicados a analizar la interpretabilidad de los SBRDs lingüísticos(ver Figura I.2):

- C_1: Complejidad a nivel de BR
- C_2: Complejidad a nivel de Partición Difusa (o BD).
- C_3: Interpretabilidad Semántica a nivel de BR.
- C_4: Interpretabilidad Semántica a nivel de Partición Difusa.

Tabla I.2: Taxonomía para analizar la interpretabilidad de los SBRDs lingüísticos

	Nivel de Base de Reglas	Nivel de Partición Difusa
Complejidad	C_1 Número de reglas Número de condiciones	C_2 Número de funciones de pertenencia Número de características
Interpretabilidad Semántica	C_3 Consistencia de las reglas Reglas disparadas al mismo tiempo Transparencia de la estructura de regla	C_4 Completitud Normalidad Distinguibilidad Complementariedad

Después de estudiar los distintos trabajos en el mencionado tema, podemos afirmar que no existe una medida única y completa para cuantificar la interpretabilidad de los modelos lingüísticos. En nuestra opinión, para conseguir una buena medida global sería necesario considerar y/o combinar las medidas adecuadas de todos los cuadrantes, a fin de tener en cuenta de forma conjunta las diferentes propiedades de interpretabilidad necesarias para este tipo de sistemas.

Las distintas medidas de cada cuadrante podrían optimizarse como distintos objetivos dentro de un marco multi-objetivo. Ésto permitiría encontrar soluciones de compromiso entre las distintas medidas teniendo en cuenta como medida indispensable la precisión del modelo. El principal problema es que hoy en día los algoritmos de optimización multi-objetivo no son capaces de manejar mucho más que tres objetivos de manera adecuada. Por lo tanto, también es necesario encontrar una manera de combinarlos en un único índice utilizando pesos o usando los operadores de agregación adecuada a fin de dar la debida importancia a una u otra medida. Una posibilidad es agregar las medidas basadas en la complejidad y las medidas basadas en semántica por separado, dando lugar a dos diferentes índices. Esto permitiría encontrar los distintos equilibrios entre la precisión, la complejidad y la interpretabilidad semántica.

Por último, debemos señalar que es necesario establecer las medidas de los distintos cuadrantes y, con respecto a la agregación de las diferentes medidas en un índice global, la forma de combinar las medidas mediante la selección de los operadores de agregación apropiada no es trivial aunque si una tarea esencial.

En este trabajo se estudia qué medidas existen en la literatura para SBRDs lingüísticos y además ayuda a decidir que medida es más adecuada según la parte del sistema en la que se aplica. El artículo asociado a esta parte es:

- M.J. Gacto, R. Alcalá, F. Herrera, Interpretability of Linguistic Fuzzy Rule-Based Systems: An Overview on Interpretability Measures. Information Sciences. (Sometido)

2.3. Algoritmos Evolutivos Multi-Objetivo que Combinan las Técnicas de Ajuste y de Selección de Reglas para Obtener Sistemas Basados en Reglas Difusas Lingüísticos Precisos y Compactos

Esta parte está compuesta de dos trabajos en los que se proponen AEMOs avanzados para conseguir el mejor equilibrio entre complejidad y precisión, es conocido que encontrar ese equilibrio no es un problema fácil. Debido a que ambos requisitos (complejidad y precisión) son contradictorios, el uso de los AEMOs permite obtener un conjunto de soluciones con distintos grados de equilibrio. En ambos algoritmos la búsqueda está centrada en precisión, debido a que es el objetivo más difícil de conseguir. Se comprobó que soluciones muy sencillas pero muy poco precisas carecían de interés, por este motivo los AEMOs propuestos focalizan la búsqueda en las soluciones más precisas, ya que los AEMOs estandard tendían a obtener un frente de Pareto suboptimal centrado en soluciones demasiado simples.

1. Propuesta Inicial de un Algoritmo Evolutivo Multi-Objetivo para Problemas de Regresión.

   Este trabajo propone la aplicación de AEMOs para obtener SBRDs con un mejor equilibrio entre interpretabilidad y precisión en problemas de modelado difuso lingüístico. Para ello, se presenta un nuevo método de post-procesamiento que considera la selección de las reglas junto con el ajuste de las funciones de pertenencia, permitiendo obtener soluciones sólo en la zona del frente de Pareto con la mayor precisión, es decir, las soluciones que contengan el menor número de reglas posible, pero aún presentando una alta precisión. Este método se basa en el algoritmo SPEA2, aplicando los operadores genéticos adecuados y algunas modificaciones que permiten concentrar la búsqueda en la zona deseada del frente de Pareto.

2. Propuesta de un Algoritmo Evolutivo Multi-Objetivo Avanzado para Problemas de Regresión: Estudio sobre Distintas Alternativas.

   En este trabajo, se extiende el algoritmo de la propuesta inicial incluyendo un operador de cruce inteligente y se estudian nuevos algoritmos de la literatura especializada desarrollados específicamente para centrar la búsqueda en la zona del frente de Pareto con el mejor equilibrio. Se propone la aplicación de AEMOs para obtener SBRDs con un buen equilibrio entre la interpretabilidad y precisión. Para ello, consideramos cuatro AEMOs diferentes que consideran dos objetivos (número de reglas y el error del sistema) y que realizan una selección de reglas y un ajuste de las parámetros para obtener soluciones con un buen equilibrio, es decir, las soluciones con el menor número de reglas posibles, pero aún presentando una buena precisión. Estos métodos aplican los operadores genéticos apropiados y se basan los algoritmos SPEA2 [ZLT01], NSGA-II [DPAM02] y dos versiones de NSGA-II para encontrar los codos del frente de Pareto [BDDO04].

Los artículos asociados a ambas partes son:

- R. Alcalá, M.J. Gacto, F. Herrera, J. Alcalá-Fdez, A Multi-objective Genetic Algorithm for Tuning and Rule Selection to Obtain Accurate and Compact Linguistic Fuzzy Rule-Based Systems. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 15:5 (2007) 539-557, doi:10.1142/S0218488507004868. Citado en 17 ocasiones.
- M.J. Gacto, R. Alcalá, F. Herrera, Adaptation and Application of Multi-Objective Evolutionary Algorithms for Rule Reduction and Parameter Tuning of Fuzzy Rule-Based Systems. Soft Computing 13:5 (2009) 419-436, doi:10.1007/s00500-008-0359-z. Citado en 11 ocasiones.

2.4. Integración de un Índice para Preservar la Interpretabilidad Semántica en la Selección y Ajuste Evolutivos Multi-Objetivo de los Sistemas Difusos Lingüísticos

En este trabajo se propone un índice de interpretabilidad semántica relativa formado por la unión de tres métricas para preservar la interpretabilidad semántica de un SBRDs mientras se realiza un ajuste de las funciones de pertenencia. Estas métricas permiten medir ciertas características de las funciones de pertenencia ajustadas respecto a las funciones de pertenencia originales. Dicho índice y las métricas están propuestos para funciones de pertenencia triangulares, pero se pueden extender con unos pequeños cambios en la formulación a gaussianas o trapezoidales. Las métricas propuestas son:

- Desplazamiento de las funciones de pertenencia (δ): Mide la proximidad de los puntos centrales de las funciones de pertenencia respecto a las posiciones originales.
- Ratio de amplitud lateral de las funciones de pertenencia (γ): Mide la diferencia del ratio izquierda/derecha del soporte de las funciones de pertenencia ajustadas respecto a las funciones de pertenencia originales.
- Similaridad del área de las funciones de pertenencia (ρ): Mide la similaridad del área de las funciones de pertenencia ajustadas respecto a las funciones de pertenencia originales.

Las definiciones de estas métricas se pueden encontrar en el artículo [GAH10] incluido como parte de esta memoria en la sección 4. Todas estas métricas son relativas, toman valores entre 0 y

1 (siendo mejores cuanto mayor sea el valor obtenido) y se agregan usando la media geométrica en un índice global llamado $GM3M$. Este índice de interpretabilidad semántica también toma valores entre 0 (menor nivel de interpretabilidad) y 1 (mayor nivel de interpretabilidad) y es definido de la siguiente forma:

$$GM3M = \sqrt[3]{\delta \cdot \gamma \cdot \rho}$$

El índice $GM3M$ puede ayudar a preservar la interpretabilidad semántica de los SBRDs lingüísticos si se utiliza como medida a maximizar. De esta forma, se puede dedicar a mantener la forma original de las funciones de pertenencia, mientras que un ajuste (o cualquier tipo de aprendizaje o mejora) de la definición de sus parámetros se lleva a cabo, y representa una medida relativa de la calidad de las particiones difusas lingüísticas, una vez que sabemos cómo deben ser las más interpretables. El índice ($GM3M$) puede trabajar con particiones difusas fuertes (que cumplen con la mayoría de las medidas absolutas propuestas en la literatura) o con particiones difusas lingüísticas definidas por un experto, siendo un índice relativo capaz de cuantificar la interpretabilidad de las particiones difusas con respecto a las originales, solucionando el problema de las medidas absolutas cuando el experto tiene los conceptos claros y no encajan con las propiedades absolutas típicamente impuestas a los SBRDs lingüísticos.

Además, se propone un AEMO avanzado con prevención de incesto y reinicio automático, que mantiene resultados muy similares o incluso mejores en precisión comparados con los resultados obtenidos sin el uso del índice. El AEMO propuesto considera los siguientes objetivos: minimizar el error cuadrático medio (precisión), minimizar el número de reglas (complejidad) y/o maximizar el índice $GM3M$ (interpretabilidad semántica).

El artículo asociado a esta parte es:

- M.J. Gacto, R. Alcalá, F. Herrera, Integration of an Index to Preserve the Semantic Interpretability in the Multi-Objective Evolutionary Rule Selection and Tuning of Linguistic Fuzzy Systems. IEEE Transactions on Fuzzy Systems 18:3 (2010) 515-531, doi:10.1109/TFUZZ.2010.2041008.

3. Comentarios Finales

3.1. Breve Resumen de los Resultados Obtenidos y Conclusiones

Como acabamos de describir, hemos seguido una línea de trabajo totalmente encadenada que comienza con una introducción al uso de los SBRDs lingüísticos y la idoneidad de usar medidas de complejidad para obtener controladores con un buen comportamiento. Seguidamente se estudia qué medidas existen en la literatura para cuantificar la interpretabilidad de los SBRDs lingüísticos. Una vez conocidas las medidas de interpretabilidad existentes, hemos diseñado AEMOs específicos avanzados que permiten combinar objetivos de precisión junto con medidas de interpretabilidad. Los AEMOs permiten obtener frentes de Pareto con una gran variedad de soluciones desde las soluciones más precisas hasta las más interpretables. Por último, hemos propuesto una medida de interpretabilidad semántica relativa para mantener la interpretabilidad de las particiones difusas cuando se realiza el ajuste de las funciones de pertenencia. Además, dicha medida se ha combinado en un AEMO con medidas de complejidad.

Es importante señalar que el comportamiento de las diferentes técnicas propuestas se ha comparado con diferentes algoritmos ya propuestos en la literatura especializada para el ajuste de las funciones de pertenencia y la selección de reglas. La presente sección se dedica a resumir las lecciones aprendidas a lo largo del trabajo realizado y a destacar las conclusiones que esta memoria aporta.

3.1.1. Mejora de Controladores Difusos Obtenidos a partir de Expertos: Un Caso de Estudio sobre un Sistema de Ventilación, Calefacción y Aire Acondicionado

En este primer estudio hemos realizado un analisis experimental sobre un problema real de control de un sistema HVAC. Respecto a los resultados obtenidos en este trabajo debemos destacar las siguientes conclusiones:

- En este tipo de problemas con restricciones especiales en cuanto al número de evaluaciones permitido y con muy alta dimensionalidad, la reducción del espacio de búsqueda que proporcionan el ajuste lateral y el ajuste lateral con amplitud, permiten considerar técnicas de optimización para obtener controladores difusos más óptimos con respeto a un enfoque clásico y con mayores grados de libertad.
- En nuestra opinión, una técnica de selección reglas es necesaria cuando los controladores difusos iniciales son obtenidos por los expertos y éstos deben ser mejorados. Por lo general, una BR obtenida por expertos incluye reglas contradictorias y redundantes que deben ser eliminadas y, en todo caso, cuando esta técnica se guía por las medidas de precisión, una regla se eliminará si empeora el rendimiento del sistema.

Finalmente, se comprobó que el problema de encontrar el equilibrio entre interpretabilidad y precisión es un problema muy complejo y no sencillo. También se comprobó que forzar la obtención de soluciones más simples puede dar lugar a la obtención de soluciones más precisas.

3.1.2. Interpretabilidad de los Sistemas Basados en Reglas Difusas Lingüísticos: Una Revisión sobre Medidas de Interpretabilidad

En este trabajo se presenta una revisión de interpretabilidad de los sistemas difusos centrada en el marco de los SBRDs lingüísticos incluyendo una lista completa de los trabajos sobre el uso de técnicas o medidas que tengan en cuenta la interpretabilidad de los SBRDs lingüísticos, como parte del problema de encontrar un buen equilibrio entre interpretabilidad y precisión. Para ello, hemos propuesto una taxonomía con cuatro cuadrantes (complejidad o interpretabilidad semántica a nivel de BR o de particiones difusas) como una forma de organizar las diferentes medidas o restricciones que encontramos en la literatura para controlar la interpretabilidad en este ámbito. Hemos analizado las diferentes medidas propuestas que compiten en los distintos cuadrantes. Dado que la interpretabilidad de los SBRDs lingüísticos es aún un problema abierto, esto ayudará a los investigadores en este campo para determinar la medida más apropiada de acuerdo con la parte de la BC en la que quieren mantener / mejorar la interpretabilidad.

Después de estudiar los diferentes trabajos en el tema, podemos afirmar que no existe una medida única y completa para cuantificar la interpretabilidad de los modelos lingüísticos. En nuestra opinión, para conseguir una buena medida global sería necesario considerar las medidas adecuadas de todos los cuadrantes, a fin de tener en cuenta de forma conjunta las diferentes propiedades requeridas en este tipo de sistemas para asegurar la interpretabilidad. Las distintas medidas, de cada cuadrante, podrían optimizarse como objetivos diferentes dentro de un AEMO o pueden ser combinadas en un sólo índice usando pesos o usando operadores de agregación adecuada a fin de dar la debida importancia a uno u otra medida. En este sentido, una de las alternativas más prometedoras es considerar las medidas de complejidad y de interpretabilidad semántica junto con la precisión dentro de un marco multi-objetivo que permita obtener los distintos grados de equilibrio entre dichas medidas.

3.1.3. Algoritmos Evolutivos Multi-Objetivo que Combinan las Técnicas de Ajuste y de Selección de Reglas para Obtener Sistemas Basados en Reglas Difusas Lingüísticos Precisos y Compactos

En este apartado del estudio hemos utilizado varios modelos diferentes de AEMOs combinando técnicas de ajuste clásico de las funciones de pertenencia junto con métodos de selección de reglas. Los AEMOs propuestos consideran los siguientes objetivos: minimizar el número de reglas (complejidad) y minimizar el error cuadrático medio (precisión)

1. Propuesta Inicial de un Algoritmo Evolutivo Multi-Objetivo para Problemas de Regresión.

   En este trabajo se propone un AEMO que presenta un buen equilibrio entre interpretabilidad y precisión en comparación con los restantes métodos analizados en el trabajo. El algoritmo propuesto ha obtenido modelos incluso con una mejor precisión que los métodos obtenidos sólo guiados por medidas de precisión. De esta manera, los resultados obtenidos han demostrado que el uso de AEMOs puede representar una forma de obtener modelos lingüísticos aún más precisos y al mismo tiempo más sencillos que los obtenidos sólo por las medidas de la precisión.

   Por otro lado, el algoritmo propuesto (SPEA2$_{ACC}$) podría ser de interés para problemas en los que a pesar de presentar un carácter multi-objetivo, necesitamos como solución no toda la frontera de Pareto, sino sólo una zona específica o de interés del mismo.

2. Propuesta de un Algoritmo Evolutivo Multi-Objetivo Avanzado para Problemas de Regresión: Estudio sobre Distintas Alternativas.

   En este trabajo, hemos analizado el uso de diferentes AEMOs: TS-SPEA2 que hace uso de SPEA2 [ZLT01], TS-NSGA-II que hace uso de NSGA-II [DPAM02], dos versiones de NSGA-II para encontrar los codos del frente de Pareto que hacen uso de los algoritmos propuestos en [BDDO04] (TS-NSGA-II$_A$ y TS-NSGA-II$_U$) y los AEMOs propuestos $TS-SPEA2_{ACC}$ y $TS-SPEA2_{ACC^2}$, para mejorar el equilibrio entre interpretabilidad y precisión de los SBRDs. Se analiza su aplicación en dos casos de estudio para obtener modelos difusos lingüísticos más simples sin perder precisión.

   De este estudio podemos destacar los siguientes puntos:

   - La mayor parte de las contribuciones en este tema se hicieron en el marco de la clasificación difusa, teniendo en cuenta SBRDs Mamdani.
   - La mayoría de los trabajos sólo tienen en cuenta las medidas cuantitativas de la complejidad del sistema para determinar la interpretabilidad de los SBRDs.
   - Ninguna de los trabajos consideran un aprendizaje o un ajuste de las funciones de pertenencia, sólo realiza un aprendizaje o una selección de reglas.
   - Los AEMOs considerados fueron ligeras modificaciones de AEMOs propuesto para uso general (MOGA, NSGA-II, etc.) o específicamente desarrollado para este problema concreto y difícil. Ésto es debido a la especial naturaleza de este problema, en el que el objetivo de la precisión es más difícil que la simplificación de los modelos difusos, por la que el frente de Pareto obtenido es sub-optimo respecto al objetivo de precisión.

Los resultados obtenidos han demostrado que el diseño específico de AEMOs puede representar una forma de obtener modelos lingüísticos aún más precisos y sencillos que los obtenidos sólo mediante medidas de precisión o mediante AEMOs estandard. En este caso (al realizar también un ajuste de los parámetros), el problema del cruce de soluciones muy diferentes, con diferente número de reglas y parámetros muy diferentes se hace más importante, ya que obtener un frente de Pareto amplio con las mejores soluciones es prácticamente imposible. De esta forma, los algoritmos específicos, $TS-SPEA2_{ACC}$ y $TS-SPEA2_{ACC^2}$, presentan los mejores resultados en terminos de precisión y complejidad. Todos los algoritmos de uso general considerados en el estudio pueden considerarse como igualmente válidos, aunque TS-SPEA2 parece tener una ligera ventaja. Dos versiones diferentes de TS-NSGA-II también se han considerado a fin de enfocar la búsqueda en las zonas más prometedoras del frente de Pareto (TS-NSGA-II$_A$ y TS-NSGA-II$_U$). Sin embargo, a pesar de que tienen un comportamiento cercano a TS-SPEA2 no muestran cambios muy significativos.

3.1.4. Integración de un Índice para Preservar la Interpretabilidad Semántica en la Selección y Ajuste Evolutivos Multi-Objetivo de los Sistemas Difusos Lingüísticos

En este trabajo, hemos propuesto un índice que ayuda a preservar la interpretabilidad semántica de los sistemas difusos lingüísticos, llamado GM3M. Este índice se dedicada a preservar la forma original de las funciones de pertenencia mientras que un ajuste de sus parámetros se lleva a cabo, y representa una medida de la calidad de la BD. Se basa en el supuesto de que la BD inicial se compone de las funciones de pertenencia adecuadas con un significado lingüístico asociado (generalmente dado por un experto). Para ello, hemos propuesto un algoritmo llamado TS_{SP2-SI}, que es un AEMO de post-procesamiento diseñado para generar un conjunto de SBRDs con diferentes soluciones de compromiso entre la precisión, la complejidad e interpretabilidad semántica. Tres criterios se han considerado: error cuadrado medio, el número de reglas y el índice GM3M propuesto.

Hemos demostrado que el uso del índice GM3M dentro de un marco evolutivo multi-objetivo ayuda a los métodos de ajuste a obtener modelos más interpretables y, al mismo tiempo más precisos. Por lo tanto, un marco multi-objetivo permite obtener SBRDs caracterizados por un mejor equilibrio entre la precisión, la complejidad e interpretabilidad semántica, respecto a los obtenidos considerando sólo la precisión como el único objetivo.

Hay que destacar que la interacción de la selección de reglas con el ajuste de las funciones de pertenencia permite la derivación de modelos mucho más precisos que al mismo tiempo mantienen la interpretabilidad semántica en un alto grado o incluso la mejoran, presentando una sinergia positiva. La selección de reglas permite una importante reducción de la complejidad del sistema. Además, se observa que TS_{SP2-SI} supera a todos los métodos analizados en todos los conjuntos de datos en el error de test y con los mejores valores de GM3M cuando se realiza un ajuste de las funciones de pertenencia. De esta manera, soluciones muy interesantes también se han obtenido con mayor precisión y niveles muy altos de interpretabilidad semántica (cerca del modelo inicial). Los resultados obtenidos muestran la utilidad del índice propuesto.

3.2. Perspectivas Futuras

A continuación, se presentan algunas líneas de trabajo futuras que se plantean a partir de los métodos propuestos en esta memoria.

1. Utilizar las distintas medidas de interpretabilidad propuestas a lo largo de esta memoria en métodos de aprendizaje de SBRDs. Además, proponer nuevos algoritmos específicos que integren estas medidas para un espacio de búsqueda aún más complicado.

2. Elección de operadores de agregación de las medidas de interpretabilidad para cada uno de los cuadrantes de la taxonomía.

   La agregación de las diferentes medidas de cada uno de los cuadrantes de la taxonomía en un único índice no es una tarea trivial. Esto es debido a que algunas medidas son subjetivas y dependen de las preferencias del usuario que dan pesos para dar más o menos importancia a cada una de las medidas.

3. Buscar/fijar nuevas medidas de interpretabilidad semántica a nivel de BR y estudiar la forma de combinarlas con las medidas de otros cuadrantes de la taxonomía, posiblemente con las medidas de interpretabilidad semántica a nivel de particiones difusas o BD.

4. Desarrollar un software para comparar modelos desde el punto de vista de la interpretabilidad. Implementar un software para la representación gráfica de los modelos obtenidos que permita estudiarlos y compararlos desde el punto de vista de la interpretabilidad mostrando en qué grado cumplen con las distintas medidas de calidad que se propongan, la forma final de las funciones de pertenencia, etc.

Parte II. Publicaciones: Trabajos Publicados, Aceptados y Sometidos

1. Mejora de Controladores Difusos Obtenidos a partir de Expertos: Un Caso de Estudio sobre un Sistema de Ventilación, Calefacción y Aire Acondicionado - *Improving Fuzzy Logic Controllers Obtained by Experts: A Case Study in HVAC Systems*

Las publicaciones en revista asociadas a esta parte son:

- R. Alcalá, J. Alcalá-Fdez, M.J. Gacto, F. Herrera, Improving Fuzzy Logic Controllers Obtained by Experts: A Case Study in HVAC Systems. Applied Intelligence 31:1 (2009) 15-30, doi:10.1007/s10489-007-0107-6.
  - Estado: Publicado
  - Índice de Impacto (JCR 2009): 0,988
  - Área de Conocimiento: Computer Science, Artificial Intelligence. Ranking 71 / 102.
  - Citas: 2

Appl Intell (2009) 31: 15–30
DOI 10.1007/s10489-007-0107-6

Improving fuzzy logic controllers obtained by experts: a case study in HVAC systems

Rafael Alcalá · Jesús Alcalá-Fdez · María José Gacto · Francisco Herrera

Published online: 28 December 2007

Abstract One important Artificial Intelligence tool for automatic control is the use of fuzzy logic controllers, which are fuzzy rule-based systems comprising expert knowledge in form of linguistic rules. These rules are usually constructed by an expert in the field of interest who can link the facts with the conclusions. However, this way to work sometimes fails to obtain an optimal behaviour. To solve this problem, within the framework of Machine Learning, some Artificial Intelligence techniques could be successfully applied to enhance the controller behaviour.

Rule selection methods directly obtain a subset of rules from a given fuzzy rule set, removing inefficient and redundant rules and, thereby, enhancing the controller interpretability, robustness, flexibility and control capability. Besides, different parameter optimization techniques could be applied to improve the system accuracy by inducing a better cooperation among the rules composing the final rule base.

This work presents a study of how two new tuning approaches can be applied to improve FLCs obtained from the expert's experience in non trivial problems. Additionally, we analyze the positive synergy between rule selection and tuning techniques as a way to enhance the capability of these methods to obtain more accurate and compact FLCs. Finally, in order to show the good performance of these approaches, we solve a real-world problem for the control of a heating, ventilating and air conditioning system.

Keywords HVAC systems · Fuzzy logic controllers · Genetic tuning · Linguistic 2-tuples representation · Linguistic 3-tuples representation · Rule selection

1 Introduction

One important Artificial Intelligence tool for automatic control is the use of Fuzzy Logic Controllers (FLCs). FLCs are Fuzzy Rule-Based Systems (FRBSs) comprising expert knowledge in form of linguistic rules. Frequently, these rules are constructed by experts in the field of interest who can link the facts or evidence with the conclusions. In case of simple problems, an expert should have no problems in obtaining appropriate rules presenting a good cooperation. However, in the case of real complex problems with many variables and rules, this way to work fails to obtain an optimal performance as it is very difficult for human beings to ensure a good cooperation among rules.

To solve this problem, within the framework of Machine Learning, some Artificial Intelligence techniques could be successfully applied to enhance the controller performance. One of the most widely-used approaches for improving the performance of FRBSs, known as *tuning*, consists of refining a previous definition of the Data Base (DB) once the Rule Base (RB) has been obtained [1, 7, 11, 17, 23, 24] (in our case by experts). Classically, the tuning methods refine the three definition parameters that identify triangular Membership Functions (MFs) associated to the labels comprising

R. Alcalá (✉) · J. Alcalá-Fdez · M.J. Gacto · F. Herrera
Department of Computer Science and Artificial Intelligence,
University of Granada, 18071 Granada, Spain
e-mail: alcala@decsai.ugr.es

J. Alcalá-Fdez
e-mail: jalcala@decsai.ugr.es

M.J. Gacto
e-mail: mjgacto@ugr.es

F. Herrera
e-mail: herrera@decsai.ugr.es

the DB [11, 12] in order to find its best global configuration (to induce to the best cooperation among the rules). However, in the case of problems with many variables, the dependency among MFs and the dependency among the three definition points, leads to tuning models handling very complex search spaces which affect the good performance of the optimization methods [4].

Recently, two new linguistic rule representation models have been proposed in order to face this particular problem [4, 5]:

- The first one was proposed to perform a genetic lateral tuning of MFs [4]. This new approach is based on the linguistic 2-tuples representation [16] that allows the symbolic translation of a label by only considering one parameter per label and therefore involves a reduction of the search space that eases the derivation of optimal models with respect to the classic tuning.
- The second one was presented to perform a fine genetic Lateral and Amplitude tuning (LA-tuning) of MFs [5]. This is based on the linguistic 3-tuples approach [5] by proposing a new symbolic representation with three values (s, α, β), respectively representing a label, the lateral displacement and the amplitude variation of the support of this label. Tuning of both parameters also involves a reduction of the search space that eases the derivation of optimal models with respect to the classic tuning.

In addition, rule selection methods directly obtain a subset of rules from a given fuzzy rule set, removing inefficient and redundant rules and, thereby, enhancing the controller interpretability, robustness, flexibility and control capability [10, 21, 22, 25, 26]. In this way, the combination of tuning techniques with rule selection methods can present a positive synergy, reducing the tuning search space, easing the system readability and even improving the system accuracy.

In this work, we present a study of how these new tuning approaches can be applied to improve FLCs obtained from the expert's experience in non trivial problems. Additionally, we analyze the positive synergy between rule selection and tuning techniques as a way to enhance the capability of these methods to obtain more accurate and compact FLCs. To show the good performance of these approaches we solve a real-world problem in the control of a Heating, Ventilating and Air Conditioning (HVAC) system [1], in which the initial FLC is obtained by experts.

This paper is arranged as follows. The next section presents the lateral tuning, the linguistic rule representation model (based on the linguistic 2-tuples) and details the evolutionary method proposed to perform the lateral tuning of FLCs. Section 3 presents the LA-tuning, the linguistic rule representation model (based on the linguistic 3-tuples) and describes the evolutionary algorithm to perform the LA-tuning. In Sect. 4, the cooperation between each tuning approach and a rule selection mechanism is analysed, presenting the evolutionary methods to perform them together. Section 5 presents a case study in a HVAC system control problem, establishing the objective function and describing the initial FLC variables and structure. Section 6 shows an experimental study of the methods behaviour applied to that problem. Finally, Sect. 7 points out some conclusions.

2 Lateral tuning of fuzzy logic controllers

This section introduces the lateral tuning of fuzzy systems, presenting the new structure of fuzzy rule and a global semantics-based tuning approach. Next, an evolutionary post-processing method to perform lateral tuning of FLCs obtained by experts is described. This method is based on that proposed in [4] for the global lateral tuning of FRBSs.

2.1 Linguistic 2-tuples re-presented rule and lateral tuning

In [4], a new model of tuning of FRBSs was proposed considering the linguistic 2-tuples representation scheme introduced in [16], that allows the lateral displacement of the support of a label maintaining the interpretability associated with the final linguistic model at a reasonable level. This new tuning approach was based on a simple data-driven learning method and a Genetic Algorithm (GA) guided by example data and considering a generational approach.

In [16], the lateral displacement represented by a linguistic 2-tuple is named symbolic translation of a linguistic label. The symbolic translation of a label is a number within the interval $[-0.5, 0.5)$, expressing this interval the domain of a label when it is moving between its two adjacent lateral labels (see Fig. 1a). Let us consider a set of labels S representing a fuzzy partition. Formally, to represent the symbolic translation of a label in S we have the 2-tuple,

$$(s_i, \alpha_i), \quad s_i \in S, \ \alpha_i \in [-0.5, 0.5).$$

In fact, the symbolic translation of a label involves the lateral displacement of its associated MF. As an example, Fig. 1 shows the symbolic translation of a label represented by the pair $(s_2, -0.3)$ together with the lateral displacement of the corresponding MF. Both the linguistic 2-tuples representation model and the elements needed for linguistic information comparison and aggregation, are presented and applied to the Decision Making framework in [16].

In the context of FRBSs, the linguistic 2-tuples could be used to represent the MFs comprising the linguistic rules. This way to work, introduces a new model for rule representation that allows the tuning of the MFs by learning their

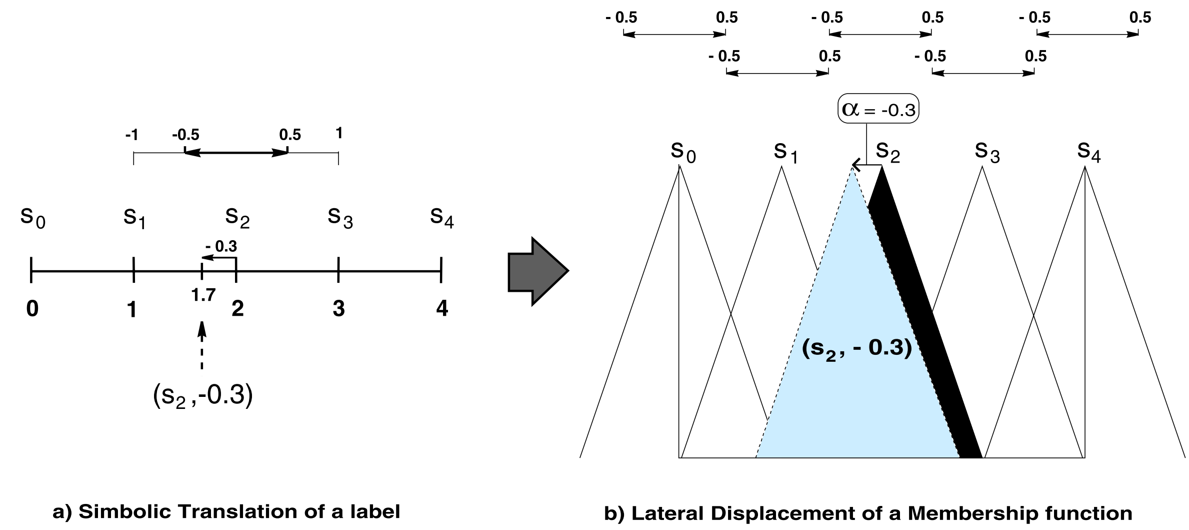


Fig. 1 Symbolic translation of a linguistic label and lateral displacement of the involved MF

respective lateral displacements. Next, we present this approach by considering a simple control problem.

Let us consider a control problem with two input variables (X1, X2), one output variable (Y) and an initial DB defined by experts to determine the MFs for the following labels:

X1: $Error \rightarrow \{Negative, Zero, Positive\}$,

X2: $\bigtriangledown Error \rightarrow \{Negative, Zero, Positive\}$,

Y: $Power \rightarrow \{Low, Medium, High\}$.

Based on this DB definition, examples of classic and linguistic 2-tuples represented rules are:

– *Classic Rule*,

R_i: If the **error** is Zero and the $\bigtriangledown$**Error** is Positive Then the **Power** is High.

– *Rule with 2-Tuples Representation*,

R_i: If the **error** is (Zero, 0.3) and the $\bigtriangledown$**Error** is (Positive, -0.2) Then the **Power** is (High, -0.1).

Analysed from point of view of rule interpretability, we could interpret the 2-tuples represented rule (i.e., a tuned rule) as:

If the **Error** is "*higher than* Zero" and
the $\bigtriangledown$**Error** is "*a little smaller than* Positive".
Then the **Power** is "*a bit smaller than* High".

In [4], two different rule representation approaches were proposed, a global approach and a local approach. In our particular case, the learning is applied to the level of linguistic partitions (global approach). In this way, the pair (X_i, label) takes the same α value in all the rules where it is considered. For example, X_i is (High, 0.3) will present the same value for those rules in which the pair "X_i is High" was initially considered. That is to say, only one displacement parameter is considered for each label on the DB.

The main difference between lateral tuning and the classic approach is the reduction of the search space focusing the search only on the MF support position, since the 3 parameters usually considered per label are reduced to only 1 symbolic translation parameter. Although lateral tuning has less freedom than the classic approach, the reduction of the search space could lead to improved performance of the tuning method, especially in complex or highly multidimensional problems, since this allows us to obtain easily the best global interaction between the MFs, thereby ensuring a good covering degree of the input data. Other important aspect is that, from the parameters α applied to each label, we could obtain the equivalent triangular MFs, by which a FRBS based on linguistic 2-tuples could be represented as a classic Mamdani FRBS [28, 29].

In this work, the fuzzy reasoning method considered is the *minimum t-norm* playing the role of implication and conjunctive operators, and the *centre of gravity weighted by the matching* strategy acts as defuzzification operator. These kinds of inference are applied once the 2-tuples represented model is transformed to (represented by) its equivalent classic Mamdani FRBS.

2.2 Algorithm for the lateral tuning

To perform the lateral tuning of MFs, in these kinds of complex problems, we consider a GA based on the well-known steady-state approach. The steady-state approach [35] consists of selecting two of the best individuals in the population and combining them to obtain two offspring. These two new individuals are included in the population replacing the two worst individuals if they are better adapted. An advantage of this technique is that good solutions are used as soon as they are available. Therefore, the convergence is accelerated while the number of evaluations needed is decreased.

In the following, the components needed to design this process are explained. They are: coding scheme and initial gene pool, chromosome evaluation, genetic operators and a restarting approach to avoid premature convergence.

– *Coding Scheme*—For the C_T part, a real coding is considered, i.e., the real parameters are the GA representation units (genes). This part is the joining of the α parameters of each fuzzy partition. Let us consider the following number of labels per variable:

  $(m^1, m^2, \ldots, m^n)$,

  with n being the number of system variables. Then, a chromosome has the form (where each gene is associated to the lateral displacement of the corresponding label in the DB),

  $C_T = (c_{11}, \ldots, c_{1m^1}, c_{21}, \ldots, c_{2m^2}, \ldots, c_{n1}, \ldots, c_{nm^n})$.

  See the C_T part of Fig. 5 (in Sect. 4) for a graphical example of coding scheme considering this approach.
– *Initial Gene Pool*—To make use of the available information, the initial FRBS obtained from expert knowledge is included in the population as an initial solution. To do so, the initial pool is obtained with the first individual having all genes with value '0.0', and the remaining individuals generated at random in $[-0.5, 0.5)$.
– *Evaluating the Chromosome*—The fitness function depends on the problem being solved (see Sect. 5.1 for our particular case of study).
– *Genetic Operators*—In part, the crossover operator is based on the concept of environments (the offspring are generated in an interval generated around their parents). These kinds of operators present good cooperation when they are introduced within evolutionary models forcing the convergence by pressure on the offspring (as in the case of the steady-state approach). Particularly, we consider a BLX-α crossover [14] and a hybrid between a BLX-α and an arithmetic crossover [18] (Fig. 2 shows the performance of these kinds of operators, which allows the offspring genes to be around a wide zone determined by both parent genes). In this way, the crossover operator is described as follows. Let us assume that $X = (x_1, \ldots, x_g)$ and $Y = (y_1, \ldots, y_g)$, $(x_i, y_i \in [a_i, b_i] \subset \Re, i = 1, \ldots, g)$, are two real-coded chromosomes that are going to be crossed:

  1. Using the BLX-α crossover [14] (with $\alpha = 0.3$), one descendent $Z = (z_1, \ldots, z_g)$ is obtained, where z_i is randomly (uniformly) generated within the interval $[l_i, u_i]$, with $l_i = \max\{a_i, c_{min} - I\}$, $u_i = \min\{b_i, c_{max} + I\}$, $c_{min} = \min\{x_i, y_i\}$, $c_{max} = \max\{x_i, y_i\}$ and $I = (c_{max} - c_{min}) \cdot \alpha$.
  2. The application of an arithmetic crossover [18] in the wider interval considered by the BLX-α, $[l_i, u_i]$, results in the next descendent:

     V with $v_i = a \cdot l_i + (1 - a) \cdot u_i$,

     where a_i and b_i are respectively -0.5 and 0.5, and $a \in [0, 1]$ is a random parameter generated each time this crossover operator is applied. In this way, this operator performs the same gradual adaptation in each gene, which involves a faster convergence in the algorithm.

  Besides, no mutation will be considered in order to favour the exploitation with respect to the exploration. For this reason, we also consider a restarting approach to avoid local optima.
– *Restarting Approach*—To get away from local optima, this algorithm uses a restart approach [13]. In this case, the best chromosome is maintained and the remaining are generated at random within the corresponding variation intervals $[-0.5, 0.5)$. It follows the principles of CHC [13], performing the restart procedure when the difference between the worst and the best chromosome fitness values is less than 1% of the initial solution fitness value. This way to work allows the algorithm to perform a better exploration of the search space and to avoid getting stuck at local optima.

Finally, the main steps of the algorithm can be found in Fig. 3 by taking into account the described components.

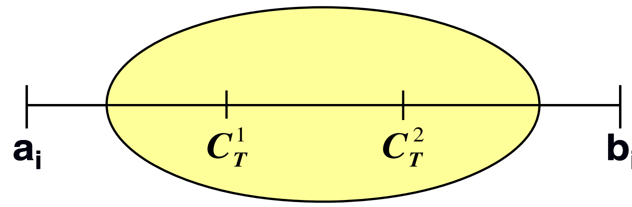


Fig. 2 Diagram of performance of the crossover operators based on environments

3 The LA-tuning of fuzzy logic controllers

This section introduces the lateral and amplitude tuning of fuzzy systems, presenting the new structure of fuzzy rule and a global semantics-based tuning approach. Then, the evolutionary post-processing method to perform LA-tuning of FLCs obtained by experts is described. This method is based on that proposed in [5] for the global LA-tuning of FRBSs.

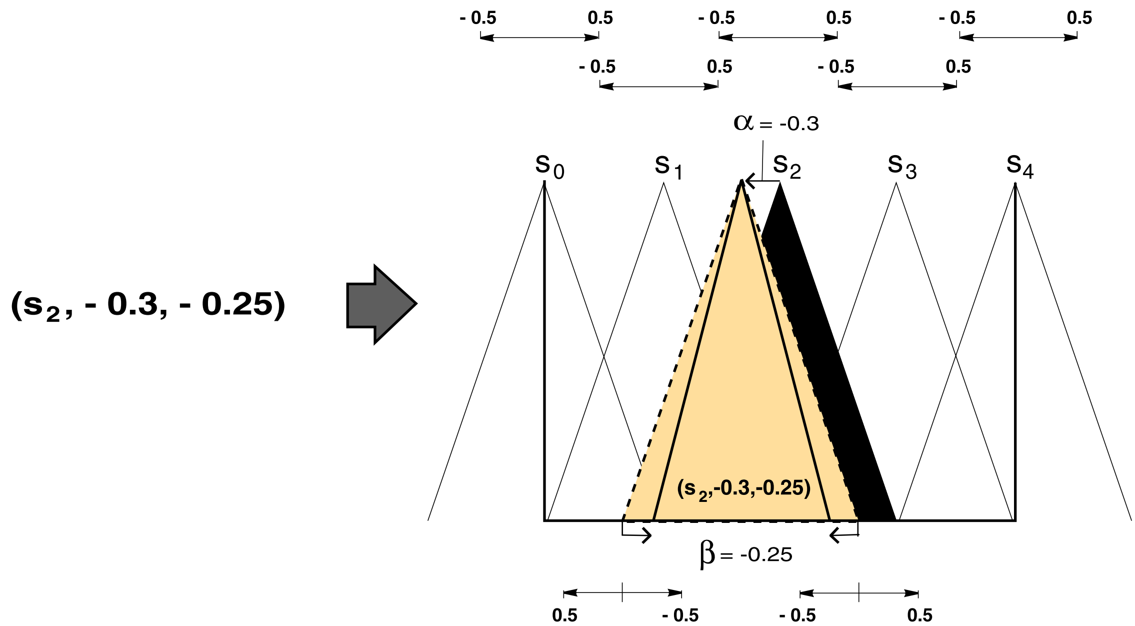


Fig. 4 Lateral and amplitude variation of the MF associated to s_2

1. Generate the initial population with N chromosomes.
2. Evaluate the population. Let F_{ini} be the tness of the initial solution obtained by experts.
3. Perform a probabilistic selection of two of the best individuals in the population.
4. Cross these individuals to obtain two offspring (hybrid BLX-α/arithmetic).
5. Evaluate the two offspring.
6. Replace the two worst individuals in the population by the two new individuals if they are better adapted. Let F_{best} and F_{worst} be the best and the worst chromosome tness values.
7. If $(F_{worst} - F_{best} < 0.01 * F_{ini})$, restart the entire population but the best.
8. If the maximum number of evaluations is not reached, go to Step 3.

Fig. 3 Scheme of the algorithm

3.1 Linguistic 3-tuples re-presented rule and LA-tuning

The LA-tuning [5] is an extension of the lateral tuning to perform also a tuning of the support amplitude of the MFs. This new approach was also based on a simple data-driven learning method and a GA guided by example data and considering a generational approach.

Determining the amplitude of a MF is a way to decide which examples are covered or not, better grouping a set of data. Therefore, tuning the amplitude of the MFs can help,

- To decrease the number of negative examples (those covered in the antecedents but not in the consequents),
- To increase the number of positive examples (those covered in the antecedents and also in the consequents), or
- To reduce the number of rules if a rule selection method is considered.

To adjust the displacements and amplitudes of the MF supports we propose a new rule representation model that considers two parameters, α and β, relatively representing the lateral displacement and the amplitude variation of a label. In this way, each label can be represented by a 3-tuple (s, α, β), where α is a number within the interval $[-0.5, 0.5)$ that expresses the domain of a label when it is moving between its two adjacent lateral labels (as in the 2-tuples representation), and β is also a number within the interval $[-0.5, 0.5)$ that allows to increase or reduce the support amplitude of a label until 50% of its original size. Let us consider a set of labels S representing a fuzzy partition. Formally, we have the triplet,

$$(s_i, \alpha_i, \beta_i), \quad s_i \in S, \ \{\alpha_i, \beta_i\} \in [-0.5, 0.5).$$

As an example, Fig. 4 shows the 3-tuple represented label $(s_2, -0.3, -0.25)$ together with the lateral displacement and amplitude variation of the corresponding MF. Let c_{s_2} and a_{s_2} be the right and the left extreme of the s_i support, and Sup_{s_2} be its size. The support of the new label $s'_2 = (s_2, -0.3, -0.25)$, can be computed in the following way:

$$Sup_{s'_2} = Sup_{s_2} + \beta * Sup_{s_2}, \quad \text{with } Sup_{s_2} = c_{s_2} - a_{s_2}.$$

In [5], two different rule representation approaches were proposed for the LA-tuning of MFs, a global approach and a local a pproach. In our case, the tuning is applied to the level of linguistic partitions (global approach). In this way, the pair (X_j, label) takes the same tuning values in all the rules where it is considered. For example, X_j is (High, 0.3, 0.1) will present the same values for those rules in which the pair "X_j is High" was initially considered. Notice that, since symmetrical triangular MFs and a FITA (*First Infer, Then Aggregate*) fuzzy inference was considered (the same presented in Sect. 2.1), a tuning of the amplitude of the consequents has no sense, by which the β parameter will be applied only on the antecedents.

In the context of FRBSs, considering the same control problem of Sect. 2.1, an example of a 3-tuples represented rule is (amplitude variation only applied in the antecedents):

R_i: If the **error** is (Zero, 0.3, 0.1) and the $\bigtriangledown$**Error** is (Positive, −0.2, −0.4). Then the **Power** is (High, −0.1).

Analised from the rule interpretability point of view, we could interpret the lateral displacement as said in Sect. 2.1. However, it is not clear a meaning for the amplitude factor β. In this way, if the final MFs are more or less well distributed and no strong amplitude changes have been performed, an expert could perhaps rename these labels giving them a more or less representative meaning. In any case, the tuning of the support amplitude keeps the shape of the MFs (triangular and symmetrical). In this way, from the parameters α and β applied to each linguistic label, we could obtain the equivalent triangular MFs, by which the last tuned FRBS could be finally represented as a classic Mamdani FRBS [28, 29].

Both approaches, lateral tuning and LA-tuning, present a good trade-off between interpretability and accuracy. However, this approach is closer to the accuracy than the lateral tuning, being this last closer to the interpretability. The choice between how interpretable and how accurate the model must be, usually depends on the user's preferences for a specific problem and it will condition the selection of the type of tuning considered (lateral or LA-tuning).

In this case, the search space increases with respect to the lateral tuning of MFs, making more difficult the derivation of optimal models. However, this approach still involves a reduction of the search space with respect to the classic tuning (one less parameter per MF), which is still well handled by means of a smart use of the search technique.

3.2 Algorithm for the LA-tuning

To perform an LA-tuning of FLCs obtained by experts we consider the same algorithm presented in Sect. 2.2 for the lateral tuning of MFs by changing the coding scheme to also consider the amplitude parameters.

In this case, the coding scheme consists in the joining of the parameters of the fuzzy partitions, lateral (C^L) and amplitude (C^A) tuning. Let us consider the following number of labels per variable: $(m^1, \dots, m^n)$, with n being the number of system variables ($n-1$ input variables and 1 output variable). Next, a chromosome has the following form (where each gene is associated to the tuning value of the corresponding label),

$$C_T = (C^L + C^A),$$
$$C^L = (c^L_{11}, \dots, c^L_{1m^1}, \dots, c^L_{n1}, \dots, c^L_{nm^n}),$$
$$C^A = (c^A_{11}, \dots, c^A_{1m^1}, \dots, c^A_{(n-1)1}, \dots, c^A_{(n-1)m^{n-1}}).$$

See the C_T part of Fig. 6 (in the next section) for a graphical example of coding scheme considering this approach.

4 Interaction between rule selection and the tuning approaches

Sometimes, a large number of fuzzy rules must be used to reach an acceptable degree of accuracy. However, an excessive number of rules makes it difficult to understand the model operation. Moreover, we may find different kinds of rules in a large fuzzy rule set: *irrelevant rules*, which do not contain significant information; *redundant rules*, whose actions are covered by other rules; *erroneous rules*, which are incorrectly defined and distort the FRBS performance; and *conflicting rules*, which perturb the FRBS performance when they coexist with others. These kinds of rules are usually obtained in non trivial problems when the final RB is generated by only considering the expert's knowledge.

To face this problem, a fuzzy rule set reduction process can be developed to achieve the goal of minimizing the number of rules used while maintaining (or even improving) the FRBS performance. To do that, erroneous and conflicting rules that degrade the performance are eliminated, obtaining a more cooperative fuzzy rule set and therefore involving a potential improvement in the system accuracy. Moreover, in many cases accuracy is not the only requirement of the model but also interpretability becomes an important aspect. Reducing the model complexity is a way to improve the system readability, i.e., a compact system with few rules requires a minor effort to be interpreted.

Fuzzy rule set reduction is generally applied as a post-processing stage, once an initial fuzzy rule set has been derived. We may distinguish between two main different approaches to obtain a more *compact* fuzzy rule set:

– *Selecting fuzzy rules*—This involves obtaining an optimal subset of fuzzy rules from a previous fuzzy rule set by selecting some of them. We may find several methods in rule selection, with different search algorithms that look for the most successful combination of fuzzy rules [10, 21, 22, 25].

  In [26], an interesting heuristic rule selection procedure is proposed where, by means of statistical measures,

a relevance factor is computed for each fuzzy rule in the linguistic FRBSs to subsequently select the most relevant ones. The philosophy of ordering the fuzzy rules with respect to an importance criterion and selecting a subset of the best seems something similar to the well-known orthogonal transfor mation-methods considered by Takagi-Sugeno-type FRBSs [33, 34].
– *Merging fuzzy rules*—This is an alternative approach that reduces the fuzzy rule set by merging the existing rules. In [27], the authors propose merging neighbouring rules, i.e., fuzzy rules where the linguistic terms used by the same variable in each rule are adjacent. Another proposal is presented in [19], where a special consideration to the merging order is made. In Takagi-Sugeno-type FRBSs, processes that simplify the fuzzy models by merging fuzzy rules have also been proposed [30–32].

These kinds of techniques for rule reduction could easily be combined with other post-processing techniques to obtain more compact and accurate FRBSs. In this way, several works have considered the selection of rules together with the tuning of MFs by coding all of them (fuzzy rules and tuning parameters) within the same chromosome [9, 15].

4.1 Positive synergy between both approaches

There are several reasons explaining the positive synergy between the rule selection and the tuning of MFs. Some of them are:

– The tuning process is affected when erroneous or conflictive rules are included in the initial RB. When the RB of a model being tuned contains bad rules (greatly increasing the system error), the tuning process tries to reduce the effect of these kinds of rules, adapting them and the remaining ones to avoid the bad performance of such rules. This way of working imposes strict restrictions, reducing the process ability to obtain precise linguistic models. Furthermore, in some cases this also affects the interpretability of the model, since the MFs comprising bad rules do not have the shape and location which best represents the information being modelled.

  This problem grows as the problem complexity grows (i.e., problems with a large number of variables and/or rules) and when the rule generation method does not ensure the generation of rules with good quality (e.g., when the initial RB is obtained by experts). In these cases, the tuning process is very complicated because the search ability is dedicated to reducing the bad performance of some rules instead of improving the performance of the remaining ones. In these cases, rule selection could help the tuning mechanism by removing the rules that really degrade the accuracy of the model.
– Sometimes redundant rules can not be removed by only using a rule selection method, since these kinds of rules could reinforce the action of poor rules improving the model accuracy. The tuning of MFs can change the performance of these rules making the reinforce ment action unnecessary, and therefore, helping the rule selection technique to remove redundant rules.

Therefore, combining rule selection and tuning approaches could cause important improvements in the system accuracy, maintaining the interpretability at an acceptable level [3, 9, 15]. However, in some cases, the search space considered when both techniques are combined is too large, which could provoke the derivation of sub-optimal models [9].

In this section, we propose the selection of a cooperative set of rules from a candidate fuzzy rule set together with the lateral or LA-tuning. This pursues the following aims:

– To improve the linguistic model accuracy selecting the set of rules best cooperating while lateral or LA-tuning is performed to improve the global configuration of MFs.
– To obtain simpler, and thus easily understandable, linguistic models by removing unnecessary or unimportant rules.
– To favour the combined action of the tuning and rule selection strategies (which involves a larger search space) by considering the simpler search space of the lateral or LA-tuning (only one or two parameters per label).

4.2 Algorithms for tuning and rule selection

To select the subset of rules which cooperate best and to obtain the tuning parameters, we consider a GA which codes all of them (rules and parameters) in one chromosome. In this way, we present two methods (one performing lateral tuning and the other performing LA-tuning) that are based on the algorithms proposed in Sects. 2.2 and 3.2, again considering the steady-state approach [35].

To do so, we must take into account the existence of binary genes (rule selection) and real values within the same chromosome. Therefore, the algorithms proposed in Sects. 2.2 and 3.2 are changed in order to consider a double coding scheme and to apply the appropriate genetic operators for each chromosome part. The following changes are considered in both algorithms in order to integrate the reduction process with the tuning of MFs:

– *Coding Scheme*—A double coding scheme for both tuning of parameters and rule selection is considered:

  $$C = C_T + C_S.$$

  In this case, the previous approaches (part C_T) are combined with the rule selection by allowing an additional binary vector C_S that directly determines when a rule is selected or not (alleles ‘1’ and ‘0’ respectively).

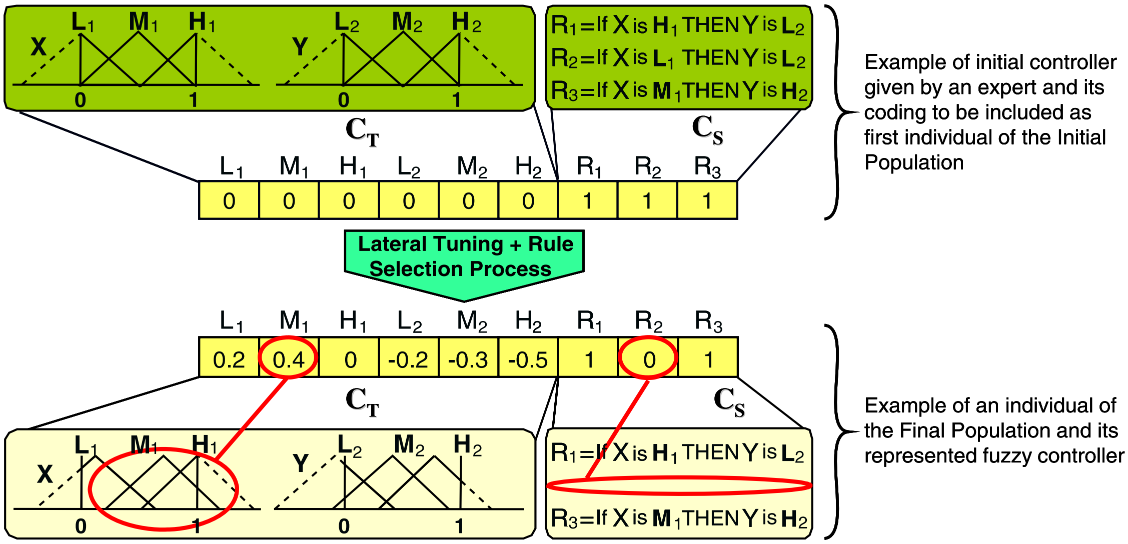


Fig. 5 Example of coding scheme considering lateral tuning and rule selection

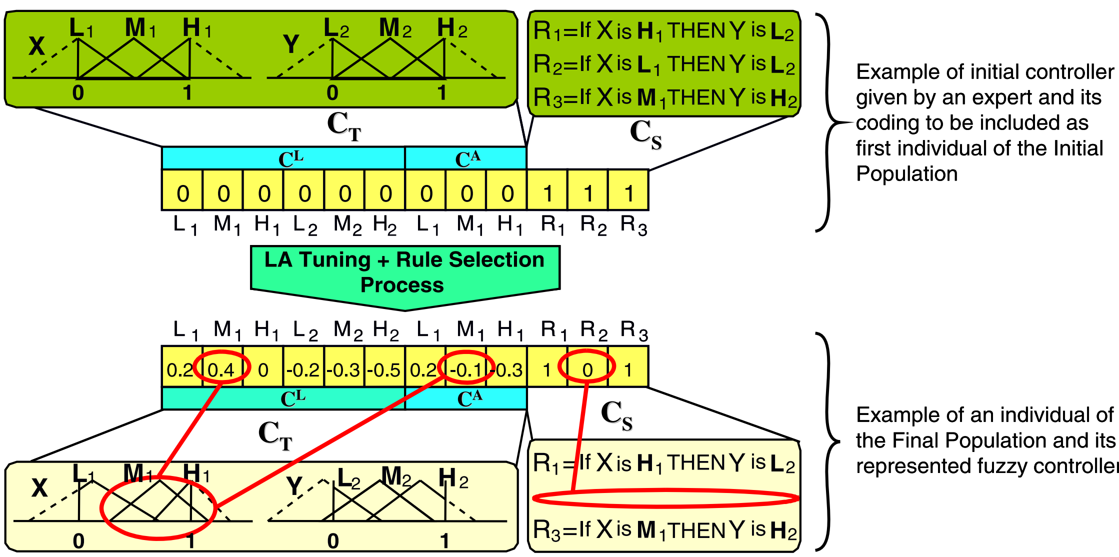


Fig. 6 Example of coding scheme considering LA-tuning and rule selection

Considering the M rules contained in the preliminary/candidate rule set, the chromosome part,

$$C_S = (c_1, \ldots, c_M),$$

represents the subset of rules composing the final rule base, such that:

$$\text{If } c_i = 1 \text{ then } (R_i \in \text{RB}) \text{ else } (R_i \not\in \text{RB}),$$

with R_i being the corresponding ith rule in the candidate rule set and RB the final rule base. Figures 5 and 6 respectively show an example of correspondence between a chromosome and its associated KB considering the lateral tuning and considering the LA-tuning.

- *Initial gene pool*—The initial pool is obtained with an individual having all genes with value '0.0' in the C_T part and '1' in the C_S part, and the remaining individuals generated at random in $[-0.5, 0.5)$ and $\{0, 1\}$ respectively.
- *Crossover*—The crossover operator presented in Sect. 2.2 for the C_T part combined with the standard two-point crossover in the C_S part. The two-point crossover operator involves exchanging the fragments of the parents contained between two points selected at random, resulting in two different offspring. In this case, four offspring are generated by combining the two from the C_T part with the two from the C_S part. The two best offspring obtained in this way are finally considered as the two co rresponding descendents.

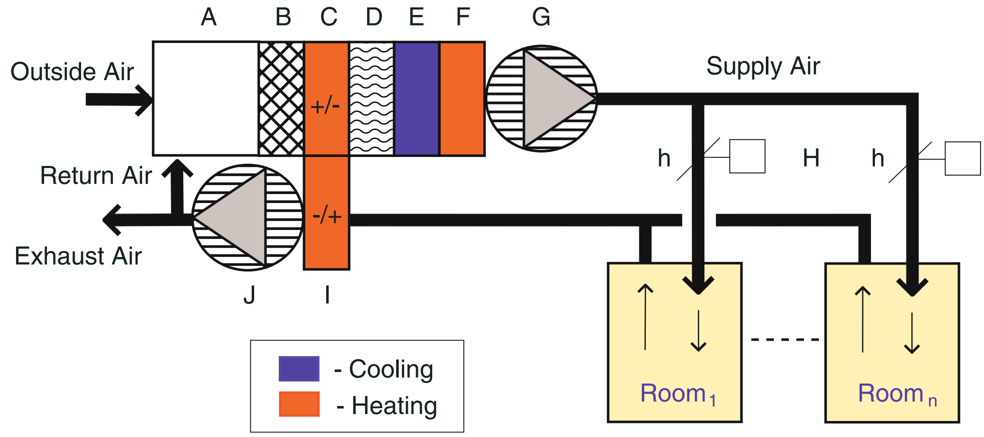


Fig. 7 Generic structure of an office building HVAC system

– *Mutation*—A mutation operator is applied on the C_S part of the four offspring before selecting the two descendents. This operator flips a gene value in C_S and helps to avoid a premature convergence in this part of the chromosome.

The application of these changes on the algorithms proposed in Sects. 2.2 and 3.2 gives rise to two different algorithms: Lateral Tuning + Rule Selection and LA − tuning + Rule Selection.

5 A case study: the HVAC system control problem

In EU countries, primary energy consumption in buildings represents about 40% of total energy consumption and more than a half of this energy is used for indoor climate conditions. On a technological point of view, it is estimated that the consideration of specific technologies like Building Energy Management Systems (BEMSs) can save up to 20% of the energy consumption of the building sector, i.e., 8% of the overall Community consumption. With this aim, BEMSs are generally applied only to the control of active systems, i.e., HVAC systems.

An HVAC system is comprised by all the components of the appliance used to condition the interior air of a building. The HVAC system is needed to provide the occupants with a comfortable and productive working environment which satisfies their physiological needs. In Fig. 7, a typical office building HVAC system is presented. This system consists of a set of components that make it possible to raise and to reduce the temperature and relative humidity of the air supply.

The energy consumption as well as indoor comfort aspects of ventilated and air conditioned buildings are highly dependent on the design, performance and control of their HVAC systems and equipments. Therefore, the use of appropriate automatic control strategies, as FLCs, for HVAC systems control could result in important energy savings when they are compared to manual control [1, 20].

Some artificial intelligence techniques could be successfully applied to enhance the HVAC system capabilities [8, 20]. However, most works apply FLCs to individually solve simple problems such as thermal regulation (maintaining the temperature at a set point), energy savings or comfort improvements. On the other hand, the initial rule set is usually constructed based on the operator's control experience using rules of thumb, which sometimes fail to obtain satisfactory results [20]. Therefore, the different involved criteria should be optimized for a good performance of the HVAC System. Usually, *the main objective is to reduce the energy consumption maintaining a desired comfort level*.

In our case, five criteria should be optimized improving an initial FLC obtained from human experience (involving 17 variables) by the application of the proposed technique for the lateral tuning of the MFs and rule selection. To do so, we consider a well calibrated and well validated model of a real test building. Both, the initial FLC and the simulation model were developed within the framework of the JOULE-THERMIE programme under the GENESYS[1] project. From now on, this test building will be called the GENESYS test site.

In the following subsections the five different objectives and the final fitness function to be optimized will be presented together with the initial FLC architecture and variables (see [1] for a more detailed information on this problem).

5.1 Objectives and fitness function

Our main optimization objective is the energy performance but maintaining the required indoor comfort levels. In this

[1] GENESYS Project: Fuzzy controllers and smart tuning techniques for energy efficiency and overall performance of HVAC systems in buildings, European Commission, Directorate-General XII for Energy (contract JOE-CT98-0090).

way, the global objective is to *minimize* the following five criteria:

O_1 Upper thermal comfort limit: if $\text{PMV} > 0.5$, $O_1 = O_1 + (\text{PMV} - 0.5)$, where PMV is the more global Predicted Mean Vote thermal comfort index 7730 selected by the international standard organization ISO, incorporating relative humidity and mean radiant temperature.[2]

O_2 Lower thermal comfort limit: if $\text{PMV} < -0.5$, $O_2 = O_2 + (-\text{PMV} - 0.5)$.

O_3 Indoor air quality requirement: if CO_2 conc. > 800 ppm, $O_3 = O_3 + (CO_2 - 800)$.

O_4 Energy consumption: $O_4 = O_4 + \text{Power at time } t$.

O_5 System stability: $O_5 = O_5 + \text{System change from time}$ t to $(t-1)$, where system change states for a change in the system operation, e.g., it counts the system operation changes (a change in the fan coil speed, extract fan speed or valve position adds 1 to the final count).

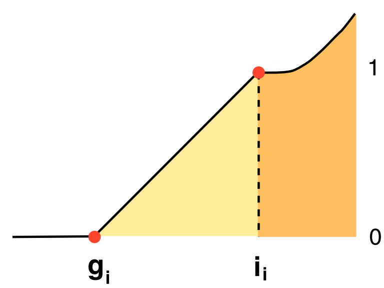
In our case, these criteria are combined into one overall objective function by means of a vector of weights. This technique (objective weighting) has much sensitivity and dependency toward weights. However, when trusted weights are available, this approach reduces the size of the search space providing the adequate direction into the solution space and its use is highly recommended. Since trusted weights were obtained from experts, we followed this approach.

Hence, an important outcome was to assign appropriate weights to each criterion of the fitness function. Although it is not part of this work and these weights were obtained within the framework of the GENESYS project, the basic idea in this weight definition was to find financial equivalents for all of them. Such equivalences are difficult to define and there is a lack of confident data on this topic. Whereas energy consumption cost is easy to set, comfort criteria are more difficult. Several studies have shown that a 18% improvement in people's satisfaction about indoor climate corresponds to a 3% productivity improvement for office workers. Based on typical salaries and due to the fact that PMV and CO_2 concentrations are related to people's satisfaction, such equivalences can be defined. The same strategy can be applied to the systems stability criterion, life-cycle of various systems being related to number of operations. Based on this, weights can be obtained for each specific building (test site). Thus, trusted weights were obtained by the experts for the objective weighting fitness function: $w_1^O = 0.0083022$, $w_2^O = 0.0083022$, $w_3^O = 0.00000456662$, $w_4^O = 0.0000017832$ and $w_5^O = 0.000761667$. Finally, the fitness function that has to be minimized was computed as:

$$F = \sum_{i=1}^{5} w_i^O \cdot O_i .$$

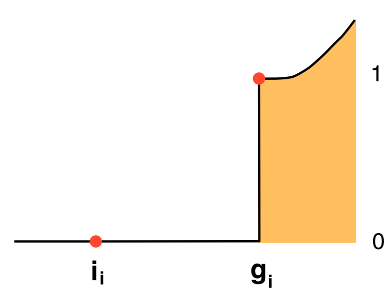
However, the fitness function has been modified in order to also consider the use of fuzzy goals that decrease the importance of each individual fitness value whenever it reaches its goal or penalize each objective whenever its value gets worse with respect to the initial solution. To do so, a function modifier parameter is considered, $\delta_i(x)$ (taking values over 1.0). A penalization rate, p_i, has been included in $\delta_i(x)$, allowing the user to set up priorities in the objectives (with 0 representing less priority and 1 more priority). Therefore, the global fitness is evaluated as:

$$F' = \sum_{i=1}^{5} w_i^O \cdot \delta_i(O_i) \cdot O_i .$$



$$\delta_i(x) = \begin{cases} 0, & \text{if } x \le g_i, \\ \dfrac{x - g_i}{i_i - g_i}, & \text{if } g_i < x < i_i, \\ \dfrac{x - i_i}{x - x \cdot p_i} + 1, & \text{if } i_i \le x. \end{cases}$$

Fig. 8 $\delta_i(x)$ when $g_i \le i_i$



$$\delta_i(x) = \begin{cases} 0, & \text{if } x < g_i, \\ \dfrac{x - g_i}{x - x \cdot p_i} + 1, & \text{if } g_i \le x. \end{cases}$$

Fig. 9 $\delta_i(x)$ when $g_i > i_i$

[2]http://www.iso.org/iso/en/ISOOnline.frontpage

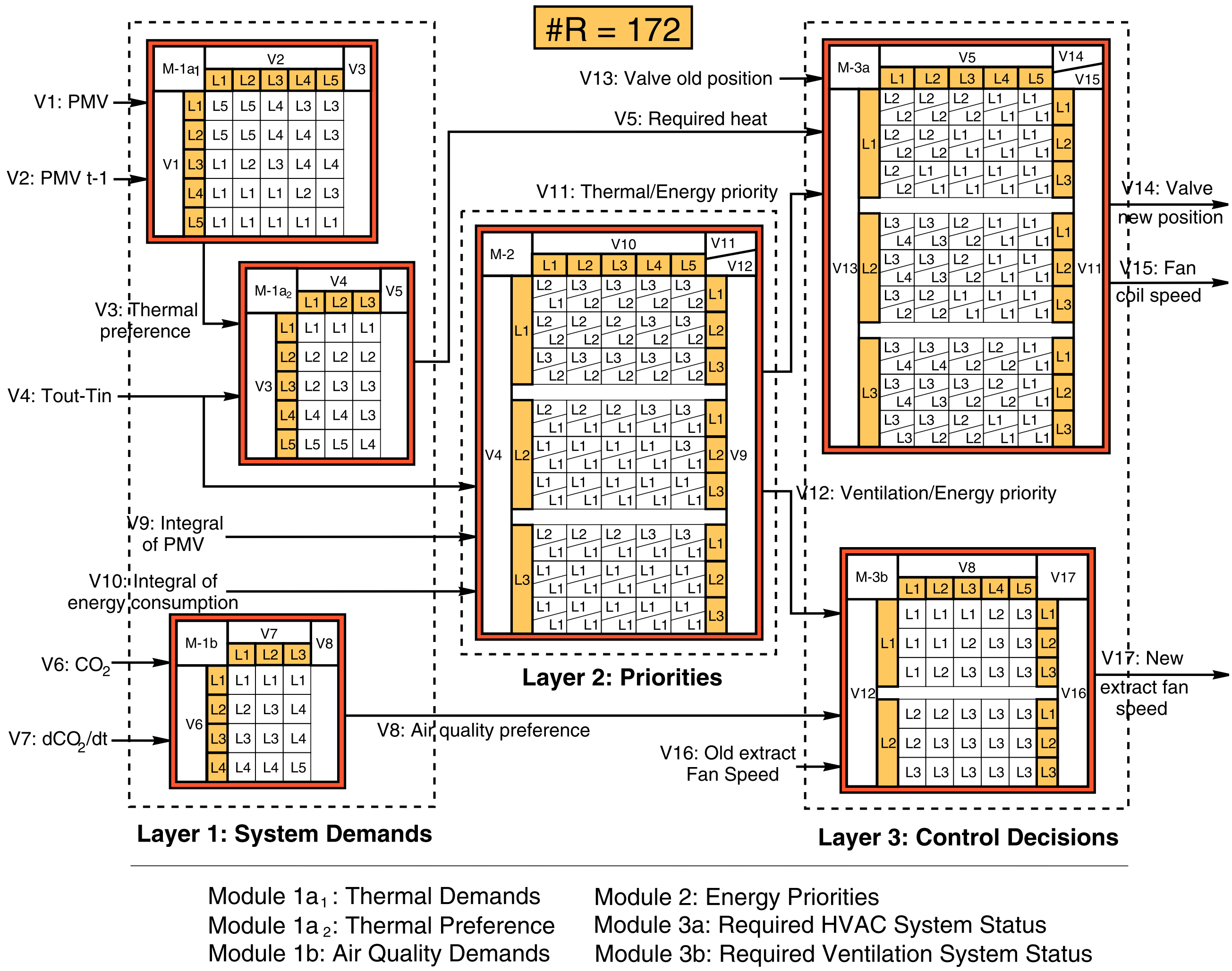


Fig. 10 Initial RB and generic structure of the GENESYS FLC

Two situations can be presented according to the value of the goal g_i, and the value of the initial solution i_i. Depending on these values, two different δ functions will be applied:

- When the value of g_i is minor than the value of i_i, the objective is not considered if the goal is met and penalized if the initial results get worse (see Fig. 8).
- When the value of i_i is minor than the value of g_i, this initial result may get worse while the goal is met and, it is penalized otherwise (see Fig. 9).

5.2 FLC variables and architecture

A hierarchical FLC architecture considering the PMV, CO_2 concentration, previous HVAC system status and outdoor temperature was proposed by the BEMS designer for this site. This architecture, variables and initial RB are presented in Fig. 10. There are three different parts (layers) in the proposed structure. The first one is devoted to the system demands, i.e., this layer analyzes the current system state and determines the required heat and the air quality preference in order to ensure a good comfort level. The second one analyzes the trend of the system in terms of PMV and energy consumption, also taking into account the outdoor and indoor temperatures in order to determine whether the system should save energy or to spend some energy to achieve a better thermal point or to perform ventilation. Finally, the third one determines the operation mode (manipulating three actuators) by taking into account the current state of the actuators and the system preferences and priorities determined by layers 1 and 2. A more detailed description of the variables considered in the initial FLC structure can be found at [1].

The DB is composed of symmetrical fuzzy partitions with triangular-shaped MFs labelled from $L1$ to Ll_i (with l_i being the number of labels of the ith variable). The initial DB is depicted in Fig. 11 together with the tuned DB. Figure 10 represents the decision tables of each module of the hierarchical FLC in terms of these labels. Each cell of the table represents a fuzzy subspace and contains its asso-

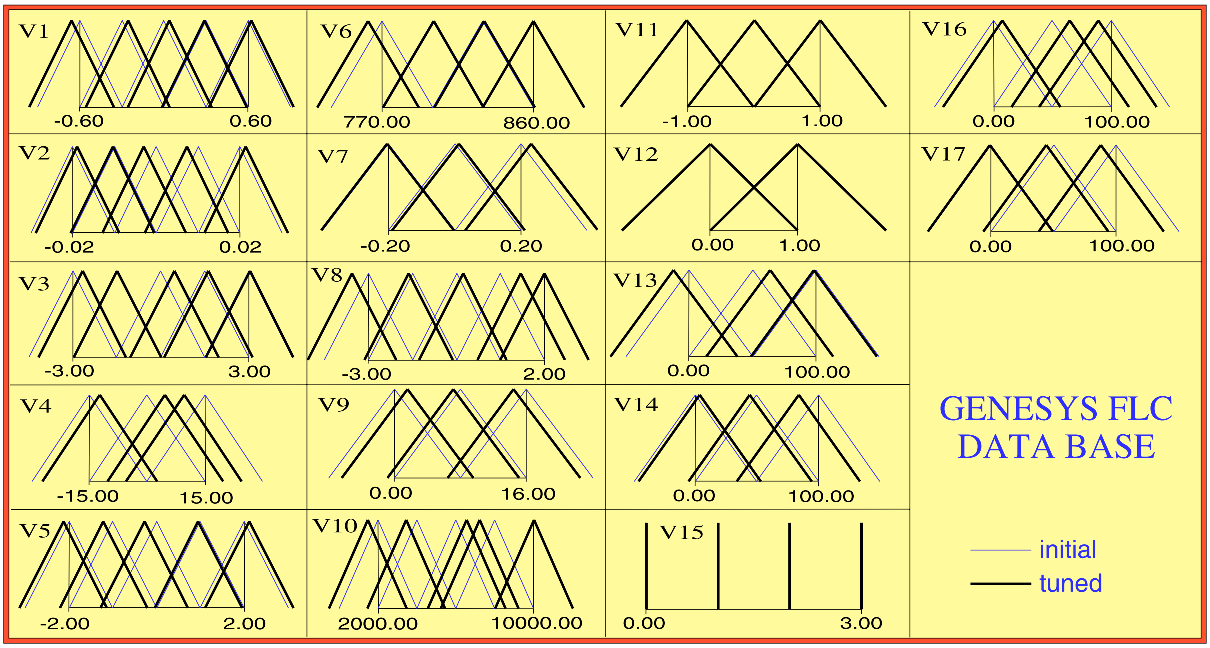


Fig. 11 Initial and tuned DB of a model obtained with GL-S (seed 1)

ciated output consequent(s), i.e., the corresponding label(s). The output variables are denoted in the top left square for each module in the figure. Both, the initial RB and the DB, were provided by the BEMS designer.

6 Experiments

To evaluate the correctness of the approaches presented in the previous sections, the HVAC problem is considered in order to be solved. The FLCs obtained from these approaches will be compared to the performance of a classic On-Off controller and to the performance of the initial FLC (provided by experts). *The goals and improvements will be computed with respect to this classic controller as done in the GENESYS project*. The experts intention was to try to have a 10% of energy saving (O_4) together with a global improvement of the system behaviour compared to On-Off control. Comfort parame ters could be slightly increased if necessary (no more than 1.0 for criteria O_1 and O_2). The methods considered in this study are shown in Table 1. S only performs rule selection (C_S part of GL-S or GLA-S) and was first used for this problem in [2] in order to be compared with a method performing rule weighting and rule selection together (although this other method, rule weighting and selection, is not comparable we can point out that the results obtained by it are so far of the results presented in this work). C performs classic tuning and was first used for this problem in [1] as a first result from the GENESYS project. C-S has been not used before in this problem and it has been developed only for comparison purposes. The remaining approaches are those presented in this paper.

The values of the parameters used in all of these experiments are presented in the following: 31 individuals, 0.2 as mutation probability per chromosome (except for GL and GLA without mutation), 0.3 for the factor α in the hybrid crossover operator and 0.35 as factor a in the max-min-arithmetic crossover in the case of C. The termination condition is to reach 2000 evaluations in all the cases, in order to perform a fair comparative study. In order to evaluate the GA good convergence, three different runs have been performed considering three different seeds for the random number generator.

The results presented in Table 2, where % stands for the improvement rate with respect to the On-Off controller for each criterion and #R for the number of fuzzy rules, correspond to averaged results obtained from the three different runs. The results obtained with the On-Off and the initial FLC controller are also included in this table. No improvement percentages have been considered in the table for $O_1 \ldots O_3$, since these objectives have always met the experts requirements (goals) and the On-Off controller presents zero values for these objectives.

A good trade-off between energy and stability was achieved for all the new models obtained considering the LA-tuning or rule selection (GL-S, GLA and GLA-S) except that considering classic tuning, with the remaining criteria for comfort and air quality within the requested levels. GL-S presents improvement rates of about 28.6% in energy and about 29.6% in stability. In the same way, GLA presents

Table 1 Methods considered for comparison

Method	Ref.	Year	Description
S	[2]	2005	Rule selection (C_S part of GL-S)
C	[1]	2003	Classic tuning
C-S	–	–	Classic tuning + rule selection
GL	–	–	Global lateral tuning
GL-S	–	–	GL + rule selection
GLA	–	–	Global LA-tuning
GLA-S	–	–	GLA + rule selection

Table 2 Comparison among the different methods

Model	#R	PMV		CO_2	Energy		Stability	
		O_1	O_2	O_3	O_4	%	O_5	%
On-Off	–	0.0	0	0	3206400	–	1136	–
Initial FLC	172	0.0	0	0	2901686	9.50	1505	−32.48
$\overline{S}$	160	0.1	0	0	2886422	9.98	1312	−15.52
$\overline{C}$	172	0.0	0	0	2586717	19.33	1081	4.84
$\overline{C-S}$	109	0.1	0	0	2536849	20.88	1057	6.98
$\overline{GL}$	172	0.9	0	0	2325093	27.49	1072	5.66
$\overline{GL-S}$	113	0.7	0	0	2287993	28.64	800	29.58
$\overline{GLA}$	172	0.9	0	0	2245812	**29.96**	797	29.84
$\overline{GLA-S}$	**104**	0.8	0	0	2253996	29.70	634	**44.19**

improvement rates of about 29.9% in energy and 29.8% in stability and GLA-S even improves the system stability up to 44.2% by only considering 100 rules approximately. Moreover, these algorithms (including GL) present a good convergence and seem to be independent of random factors.

Taking into account the differences among the results obtained by considering classic tuning (C and C-S) and those considering lateral or LA-tuning we can point out that, in complex problems (problems in which to obtain a set with cooperative rules is non trivial for an expert), the search space is too large to obtain a good global configuration of the MFs and rules. In this manner, conside ring techniques to ease the way to obtain a more global optimum can take advantage with respect to other approaches with more freedom degrees but handling too large search spaces.

Besides, we have to highlight that the best results obtained from those methods considering rule selection with much less rules indicate that there are a lot of rules that are wrong or not necessary in the initial RB provided by an expert. Probably, many of them are contradictory rules forcing the HVAC system to continuously change its way of working instead of maintaining a stable operation mode.

Figures 11 and 12 represent the initial and final DB of a FLC obtained by GL-S and GLA-S (seed 1). They show that not so strong variations in the MFs can involve important improvements. Moreover, Fig. 13 represents the corresponding decision tables of the model obtained from GLA-S with seed 1. In this case, a large number of rules have been removed from the initial FLC, obtaining much simpler models (72 rules were removed). This fact improves the system readability, and allows us to obtain more simple and accurate FLCs.

7 Concluding remarks

In this work, we propose the use of two advanced tuning techniques (lateral and LA-tuning) and their combination with rule selection to improve FLCs obtained by experts in non trivial problems. A case study for the control of HVAC systems has been considered in order to apply these new techniques. From the results obtained we can point out the following conclusions:

– In these kinds of non trivial problems, the search space reduction that lateral and LA-tuning involve allows the considered optimization technique to obtain more optimal FLCs respect with a classic approach with more freedom degrees.
– In our opinion, a rule selection technique is necessary when an initial FLC obtained by experts is considered to be improved. Usually, a RB obtained by experts includes conflicting and redundant rules that should be removed

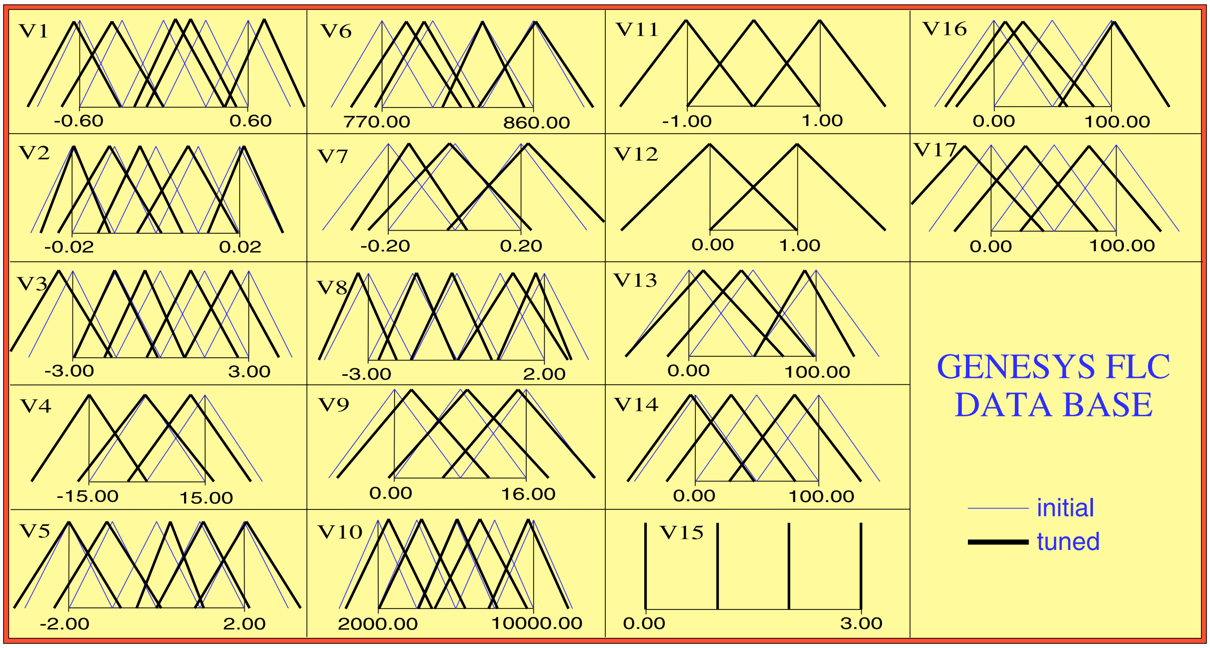


Fig. 12 Initial and tuned DB of a model obtained with GLA-S (seed 1)

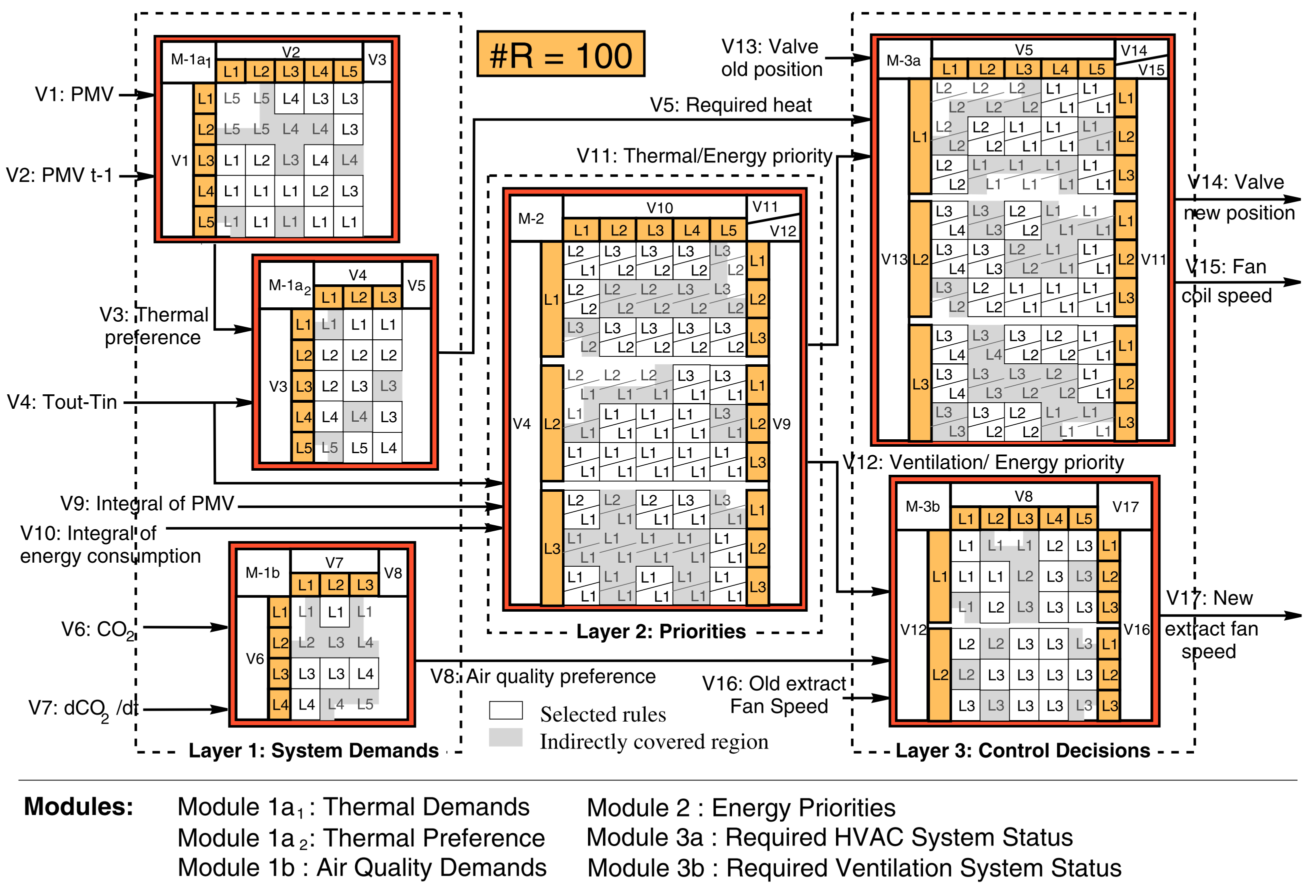


Fig. 13 RB and final structure of a model obtained with GLA-S (seed 1)

and, in any case, when this technique is guided by accuracy measures no rules will be removed if that worsen the system performance.
– The search space reduction provided by the lateral and the LA-tuning helps to better handle the larger search space that the combination between rule selection and tuning techniques involves, taking advantage respect to the classic approach.

As mentioned, tuning is a variation in the shape of the MFs that improves their global interaction with the main aim of inducing better cooperation among the rules. In this way, the real aim of the tuning is to find the best global configuration of the MFs and not only to find independently specific MFs. The main difference of lateral and LA-tuning with the classic approach is the reduction of the search space focusing the search only on the MF support position. Although lateral and LA-tuning have less freedom than the classic approach, the reduction of the search space could lead to improved performance of the tuning method, especially in complex or highly multidimensional problems, since this allows us to obtain easily the best *global interaction* between the MFs, thereby ensuring a good performance of the obtained controllers. The use of these new techniques is then justifiable when the classic approach is not able to obtain this global configuration due to the existence of a very large or complex search space. This is the case of the technique presented in [6] based on the 2-tuples representation to learn the whole knowledge base (number of MFs, rule base and parameters all together), which itself represent a very complex search space independently of the problem being solved. Unfortunately, this technique can not be applied to these kinds of problems based on an initial rule base obtained from experts since the rule base extraction is completely based on the existence of example data, and they are not usually available in these kinds of problems.

As further work, we propose the use of multiobjective GAs in order to obtain even simpler FLCs but maintaining a similar accuracy, which represent an even more complex search space.

Acknowledgements Supported by the Spanish Ministry of Education and Science under grant No. TIN2005-08386-C05-01, and the Andalusian government under grant No. P05-TIC-00531.

References

1. Alcalá R, Benítez JM, Casillas J, Cordón O, Pérez R (2003) Fuzzy control of HVAC systems optimized by genetic algorithms. Appl Intell 18:155–177
2. Alcalá R, Casillas J, Cordón O, González A, Herrera F (2005) A genetic rule weighting and selection process for fuzzy control of HVAC systems. Eng Appl Artif Intell 18(3):279–296
3. Alcalá R, Alcalá-Fdez J, Casillas J, Cordón O, Herrera F (2006) Hybrid learning models to get the interpretability-accuracy trade-off in fuzzy modeling. Soft Comput 10(9):717–734
4. Alcalá R, Alcalá-Fdez J, Herrera F (2007) A Proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection. IEEE Trans Fuzzy Syst 15(4):616–635
5. Alcalá R, Alcalá-Fdez J, Gacto MJ, Herrera F (2007) Rule base reduction and genetic tuning of fuzzy systems based on the linguistic 3-tuples representation. Soft Comput 11(5):401–419
6. Alcalá R, Alcalá-Fdez J, Herrera F, Otero J (2007) Genetic learning of accurate and compact fuzzy rule based systems based on the 2-tuples linguistic representation. International J Approx Reason 44(1):45–64
7. Babuška R, Oosterhoff J, Oudshoorn A, Bruijn PM (2002) Fuzzy self-tuning PI control of pH in fermentation. Eng Appl Artif Intell 15(1):3–15
8. Calvino F, Gennusa ML, Rizzo G, Scaccianoce G (2004) The control of indoor thermal comfort conditions: introducing a fuzzy adaptive controller. Energy Build 36:97–102
9. Casillas J, Cordón O, Del Jesus MJ, Herrera F (2005) Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction. IEEE Trans Fuzzy Syst 13(1):13–29
10. Cordón O, Herrera F (2000) A proposal for improving the accuracy of linguistic modeling. IEEE Trans Fuzzy Syst 8(3):335–344
11. Cordón O, Herrera F, Hoffmann F, Magdalena L (2001) Genetic fuzzy systems: Evolutionary tuning and learning of fuzzy knowledge bases. World Scientific, Singapore
12. Driankov D, Hellendoorn H, Reinfrank M (1993) An introduction to fuzzy control. Springer, Berlin
13. Eshelman LJ (1991) The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. In: Foundations of Genetic Algorithms, vol 1. Morgan Kaufman, San Mateo, pp 265–283
14. Eshelman LJ, Schaffer JD (1993) Real-coded genetic algorithms and interval-schemata. Found Genet Algorithms 2:187–202
15. Gómez-Skarmeta AF, Jiménez F (1999) Fuzzy modeling with hybrid systems. Fuzzy Sets Syst 104:199–208
16. Herrera F, Martńez L (2000) A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Trans Fuzzy Syst 8(6):746–752
17. Herrera F, Lozano M, Verdegay JL (1995) Tuning fuzzy logic controllers by genetic algorithms. Int J Approx Reason 12:299–315
18. Herrera F, Lozano M, Verdegay JL (1997) Fuzzy connectives based crossover operators to model genetic algorithms population diversity. Fuzzy Sets Syst 92(1):21–30
19. Hong TP, Lee CY (1999) Effect of merging order on performance of fuzzy induction. Intell Data Anal 3(2):139–151
20. Huang S, Nelson RM (1994) Rule development and adjustment strategies of a fuzzy logic controller for an HVAC system—parts I and II (Analysis and Experiment). ASHRAE Trans 100(1):841–856
21. Ishibuchi H, Murata T, Türksen IB (1997) Single-objective and two objective genetic algorithms for selecting linguistic rules for pattern classification problems. Fuzzy Sets Syst 89(2):135–150
22. Ishibuchi H, Nozaki K, Yamamoto N, Tanaka H (1995) Selecting fuzzy if-then rules for classification problems using genetic algorithms. IEEE Trans Fuzzy Syst 3(3):260–270
23. Jang JSR (1993) ANFIS: Adaptive network based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–684
24. Karr C (1991) Genetic algorithms for fuzzy controllers. AI Expert 6(2):26–33
25. Krone A, Krause H, Slawinski T (2000) A new rule reduction method for finding interpretable and small rule bases in high dimensional search spaces. Proc IEEE Int Conf Fuzzy Syst 2:693–699
26. Krone A, Taeger H (2001) Data-based fuzzy rule test for fuzzy modelling. Fuzzy Sets Syst 123(3):343–358

27. Klose A, Nurnberger A, Nauck D (1998) Some approaches to improve the interpretability of neuro-fuzzy classifiers. In: Proceedings of the 6th European congress on intelligent techniques and soft computing, pp 629–633
28. Mamdani EH (1974) Application of fuzzy algorithms for control of simple dynamic plant. Proc IEEE 121(12):1585–1588
29. Mamdani EH, Assilian S (1975) An experiment in linguistic synthesis with a fuzzy logic controller. Int J Man-Mach Stud 7:1–13
30. Setnes M, Babuška R, Kaymak U, van Nauta Lemke HR (1998) Similarity measures in fuzzy rule base simplification. IEEE Trans Syst Man Cybern B: Cybern 28(3):376–386
31. Setnes M, Roubos JA (2000) GA-fuzzy modeling and classification: complexity and performance. IEEE Trans Fuzzy Syst 8(5):509–522
32. Roubos JA, Setnes M (2001) Compact and transparent fuzzy models and classifiers through iterative complexity reduction. IEEE Trans Fuzzy Syst 9(4):516–524
33. Yam Y, Baranyi P, Yang CT (1999) Reduction of fuzzy rule base via singular value decomposition. IEEE Trans Fuzzy Syst 7:120–132
34. Yen J, Wang L (1999) Simplifying fuzzy rule-based models using orthogonal transformation methods. IEEE Trans Syst Man Cybern B: Cybern 29:13–24
35. Whitley D, Kauth J (1988) GENITOR: A different genetic algorithm. In: Proceedings of the Rocky Mountain conference on artificial intelligence, pp 118–130

Rafael Alcalá received the M.Sc. degree in Computer Science in 1998 and the Ph.D. degree in Computer Science in 2003, both from the University of Granada, Spain.
He is currently an Assistant Professor in the Department of Computer Science and Artificial Intelligence at the University of Granada, where he is a Member of the Soft Computing and Intelligent Information Systems Research Group. He has over 43 international publications, 15 of them published in international journals. As edited activities, he has co-edited the IEEE Transactions on Fuzzy Systems Special Issue on "Genetic Fuzzy Systems: What's next". He has worked on several research projects supported by the Spanish government and the European Union. His research interests include, multi-objective genetic algorithms and genetic fuzzy systems, especially the learning/tuning of fuzzy systems for modeling and control with trade-off between interpretability and accuracy.

Jesús Alcalá-Fdez received the M.Sc. and Ph.D. degrees in computer science from the University of Granada, Spain, in 2002 and 2006, respectively.
He is currently an Assistant Professor in the Department of Computer Science and Artificial Intelligence, University of Granada, where he is a Member of the Soft Computing and Intelligent Information Systems Research Group in the Department of Computer Science and Artificial Intelligence at the University of Granada. He has over 25 international publications, 12 of them published in international journals. He has worked on several research projects supported by the Spanish government.
As edited activities, he has co-edited the IEEE Transactions on Fuzzy Systems Special Issue on "Genetic Fuzzy Systems: What's next". His research interests include fuzzy association rules, data mining software, multiobjective genetic algorithms and genetic fuzzy systems, especially the learning/tuning of fuzzy systems for modeling and control with trade-off between interpretability and accuracy.

María José Gacto received the M.Sc. degree in Computer Science in 1999 from the University of Granada, Spain.
She is a Member of the Soft Computing and Intelligent Information Systems Research Group in the Department of Computer Science and Artificial Intelligence at the University of Granada. She has over 15 international publications. She has worked on several research projects supported by the Spanish government and the European Union. Her research interests include, multi-objective genetic algorithms and genetic fuzzy systems, especially the learning/tuning of fuzzy systems for modeling and control with trade-off between interpretability and accuracy.

Francisco Herrera received the M.Sc. degree in Mathematics in 1988 and the Ph.D. degree in Mathematics in 1991, both from the University of Granada, Spain.
He is currently a Professor in the Department of Computer Science and Artificial Intelligence at the University of Granada. He has published more than 100 papers in international journals. He is coauthor of the book "*Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Knowledge Bases*" (World Scientific, 2001). As edited activities, he has co-edited four international books and co-edited 16 special issues in international journals on different Soft Computing topics. He currently serves as area editor of the Journal Soft Computing (area of genetic algorithms and genetic fuzzy systems), and he serves as member of the editorial board of the journals: Fuzzy Sets and Systems, International Journal of Hybrid Intelligent Systems, International Journal of Computational Intelligence Research, Mediterranean Journal of Artificial Intelligence, International Journal of Information Technology and Intelligent and Computing, Evolutionary Intelligence, and Memetic Computation. He acts as associated editor of the journals: Mathware and Soft Computing, Advances in Fuzzy Systems, and Advances in Computational Sciences and Technology.
His current research interests include computing with words and decision making, data mining and knowledge discovery, data preparation, fuzzy rule based systems, genetic fuzzy systems, knowledge extraction based on evolutionary algorithms, memetic algorithms and genetic algorithms.

2. Interpretabilidad de los Sistemas Basados en Reglas Difusas Lingüísticos: Una Revisión sobre Medidas de Interpretabilidad - *Interpretability of Linguistic Fuzzy Rule-Based Systems: An Overview on Interpretability Measures*

Las publicaciones en revista asociadas a esta parte son:

- M.J. Gacto, R. Alcalá, F. Herrera, Interpretability of Linguistic Fuzzy Rule-Based Systems: An Overview on Interpretability Measures. Information Sciences.
  - Estado: Sometido a Revisión.
  - Índice de Impacto (JCR 2009): 3,291.
  - Área de Conocimiento: Computer Science, Information Systems. Ranking 6 / 116.

Interpretability of Linguistic Fuzzy Rule-Based Systems: An Overview on Interpretability Measures[☆]

M.J. Gacto[a], R. Alcalá[b], F. Herrera[b]

[a] *Dept. Computer Science, University of Jaén, 23071 Jaén, Spain.*
[b] *Department of Computer Science and Artificial Intelligence, University of Granada, 18071 Granada, Spain.*

Abstract

Linguistic fuzzy modelling, developed by linguistic fuzzy rule-based systems, allows us to deal with the modelling of systems by building a linguistic model which is clearly interpretable by human beings. Linguistic fuzzy modelling comes with two contradictory requirements: interpretability and accuracy. In recent years the interest of researchers in obtaining more interpretable linguistic fuzzy models has grown.

Whereas the measures of accuracy are straightforward and well-known, the interpretability measures are difficult to define since interpretability depend on several factors, mainly the model structure, the number of rules, the number of features, the number of linguistic terms, the shape of the fuzzy sets, etc. Moreover, due to the subjectivity of the concept the choice of appropriate interpretability measures is still an open problem.

In this paper, we present an overview of the proposed interpretability measures and techniques for obtaining more interpretable linguistic fuzzy rule-based systems. To this end, we will propose a taxonomy based on a double axis: "Complexity versus Semantic Interpretability" considering the two main kinds of measures; and "Rule Base versus Fuzzy Partitions" considering the different components of the knowledge base to which both kinds of measures can be applied. The main aim is to provide a well established framework in order to facilitate a better understanding of the topic and well founded future works.

Keywords: Complexity, Semantic Interpretability, Linguistic Fuzzy Rule-Based Systems

1. Introduction

Fuzzy modelling (FM), system modelling by Fuzzy Rule-Based Systems (FRBSs), may be considered as an approach used to model a system making use of a descriptive language based on fuzzy logic with fuzzy predicates. Depending on which is the main requirement pursued, the FM field may be divided into two different areas:

1. Linguistic fuzzy modelling (LFM): The main objective is to obtain fuzzy models with good interpretability and it is mainly developed by means of linguistic (or classic Mamdani) FRBSs [1, 2]. Linguistic FRBSs are based on linguistic rules, in which the antecedent and the consequent make use of linguistic variables comprised of linguistic terms and the associated fuzzy sets defining their meanings.
2. Precise fuzzy modelling (PFM): The main objective is to obtain fuzzy models with good accuracy, and it is mainly developed by means of Takagi-Sugeno FRBSs [3] or by means of approximate FRBSs, which differ from the linguistic ones in the use of fuzzy variables, i.e., fuzzy sets without an associated meaning.

In this work, we are going to focus on LFM since this approach allows us to deal with the modelling of systems by building a linguistic model which is clearly interpretable by human beings, and is therefore naturally closer to interpretability than PFM. Focusing on LFM, we have to pay attention to two contradictory requirements of the model:

[☆] Supported by the Spanish Ministry of Education and Science under grant no. TIN2008-06681-C06-01.
Email addresses: mjgacto@ugr.es (M.J. Gacto), alcala@decsai.ugr.es (R. Alcalá), herrera@decsai.ugr.es (F. Herrera)

- *Accuracy*: This is the capability to faithfully represent the real system. It should be better as there is a higher similarity between the responses of the real system and the fuzzy model. There exit well-defined measures that are widely accepted in order to assess how good the accuracy is.

- *Interpretability*: This is the capacity to express the behavior of the real system in an understandable way. It is a subjective property that depends on several factors, mainly the model structure, the number of input variables, the number of fuzzy rules, the number of linguistic terms, the shape of the fuzzy sets, etc. There is still no standard measure to assess how good interpretability is.

Researchers have usually focused on the improvement of the accuracy of the models obtained without paying special attention to the interpretability. Nowadays, the interest of the researchers in interpretability has grown, which has prompted the appearance of a great quantity of work with the purpose of obtaining more interpretable linguistic models. However, two important problems remain to be solved:

- Accuracy and interpretability represent contradictory objectives. The ideal thing would be to satisfy both criteria to a high degree, but since they are contradictory properties it is generally not possible. Because of this, researchers usually focus on obtaining the best trade-off between interpretability and accuracy [4, 5], depending on the user necessities.

- Due to its subjective nature and the large amount of factors involved, the choice of an appropriate interpretability measure is still an open problem. Most researchers would agree on interpretability involving aspects such as: the number of rules should be as small as possible; rule premises should be easy in structure and contain only a few input variables, linguistic terms should be intuitively comprehensible; etc... In this event, whereas the definition of accuracy in a certain application is straightforward, the definition of interpretability is rather problematic.

In this paper, we present an overview of the proposed interpretability measures and techniques for obtaining more interpretable linguistic FRBSs. To this end, we will propose a taxonomy based on a double axis: "Complexity versus Semantic Interpretability" considering the two main kinds of measures; and "Rule Base versus Fuzzy Partitions" considering the different components of the Knowledge Base (KB) to which both kinds of measures can be applied. This leads to four different quadrants to be analyzed: the complexity at the Rule Base (RB) level; the complexity at the fuzzy partition level; the semantic at the RB level; the semantic at the fuzzy partition level. The main aim is to provide a well established framework in order to facilitate a better understanding of the topic and well founded future works.

We consider that a revision of the existing methodologies and measures, to take into account the interpretability of linguistic FRBSs as a part of the interpretability-accuracy trade-off problem together with a taxonomy, would be very interesting in order to show how the different authors tackled this difficult problem. Although, there are two review papers by Zhou and Gan [6] and by Mencar *et al* [7] studying the interpretability of fuzzy systems in general, there is no exhaustive review of interpretability issues that specifically focuses on the area of linguistic FRBSs, which represents an extensive framework deserving a deeper analysis from the mentioned double point of view, complexity versus semantic interpretability.

This paper is arranged as follows. The next section presents the taxonomy proposed for studying interpretability in the linguistic FRBSs area. Section 3 presents the works considering the complexity at the RB level. This is one of the quadrants of the taxonomy representing the most extended way to work, i.e., the classic measures to obtain simpler models. In section 4, we present the measures that the authors proposed to improve the interpretability by taking into account the complexity at the fuzzy partition level. In section 5, we analyze those works devoted to maintaining the semantic interpretability at the RB level. Section 6 includes those works trying to ensure the semantic interpretability at the fuzzy partition level which usually imposes constraints on the MFs by considering measures such as distinguishability, coverage, etc. Finally, in section 7 we draw some conclusions.

2. Semantic Interpretability versus Complexity: A Taxonomy on Linguistic FRBSs

In this section, we propose a specific taxonomy that can help us to better understand how the interpretability aspect has been taken into account in the particular framework of linguistic FRBSs. Different works [6, 7, 8] have proposed interesting taxonomies as a way to study interpretability aspects inside the more general area of fuzzy systems:

- Alonso *et al* in [8] present a taxonomy including the main factors to be considered in order to assess the interpretability of FRBSs (for both, LFM and PFM). Those factors are description and explanation: "On the one hand, the system is viewed as a whole describing its global behavior and trend. On the other hand, each individual situation is analyzed explaining specific behaviors for specific events".
- A taxonomy on interpretability constraints for information granules has been suggested by Mencar *et al* in [7], considering: "constraints for fuzzy set, constraints for universe of discourse, constraints for fuzzy information granules, constraints for fuzzy rules, constraints for fuzzy models and constraints for learning algorithms".
- Zhou and Gan [6] propose a taxonomy in terms of low-level interpretability and high-level interpretability. *Low-level interpretability* of fuzzy models is achieved on fuzzy set level by optimizing Membership Functions (MFs) in terms of the semantic criteria on MFs (semantics-based interpretability) and *high-level interpretability* is obtained on the fuzzy rule level by conducting overall complexity reduction in terms of some criteria, such as a moderate number of variables and rules (complexity-based interpretability) and consistency of rules (semantics-based interpretability).

Following our approximation, the two main kinds of approaches in order to take into account the interpretability of linguistic FRBSs are:

1. Complexity-based Interpretability: These approaches are devoted to decreasing the complexity of the obtained model [9, 10, 11, 12, 13, 14, 15, 16, 17, 18] (usually measured as number of rules, variables, labels per rule, etc.).
2. Semantics-based Interpretability: These approaches are devoted to preserve the semantics associated to the MFs [11, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28]. We can find approaches trying to ensure semantic integrity by imposing constraints on the MFs or approaches considering measures such as distinguishability, coverage, etc.

Since both kinds of measures, complexity-based interpretability and semantic-based interpretability, should be considered in both KB components, linguistic fuzzy partition and RB, the taxonomy is based on a double axis:

- Complexity versus Semantic interpretability.
- Rule Base versus Fuzzy Partition.

Finally, the proposed taxonomy comes from combining both axes. This leads to the appearance of the following quadrants devoted to analyzing the interpretability of linguistic FRBSs (see Figure 1):

- C_1: Complexity at the RB level, analyzed in section 3.
- C_2: Complexity at the fuzzy partition level, analyzed in section 4.
- C_3: Semantic at the RB level, analyzed in section 5.
- C_4: Semantic at the fuzzy partition level, analyzed in section 6.

The *Complexity* problem has been solved in the literature in different ways. Some works use techniques such as merging [18, 29, 30, 31] (to reduce the number of MFs) or rule selection [9, 11, 12, 13, 14, 16, 17, 32] (to reduce the number or rules) or methods for rule learning [14, 17] (directly obtaining simple models). We will also consider these techniques when they explicitly mention complexity reduction as one of their inherent properties.

The following sections analyze the different quadrants trying to emphasize those interpretability criteria/measures proposed within each of them. Some works will appear in only one quadrant, while others may be in several quadrants simultaneously. There are two possible reasons to include a work within several quadrants. One reason may be that the work seeks to improve interpretability by approaching it from several angles, for example including measures like the number of rules (C_1) and distinguishability of the MFs (C_4). Another reason is that, by improving a measure of one quadrant one can also produce an improvement in the measures of a different quadrant, for example reducing the number of MFs (C_2) as a way to reduce the number of rules (C_1).

Table 1: A taxonomy to analyze the interpretability of linguistic FRBSs

	Rule Base level	Fuzzy Partition level
Complexity-based Interpretability	C_1 Number of rules Number of conditions	C_2 Number of Membership Functions Number of Features
Semantic-based Interpretability	C_3 Consistency of rules Rules fired at the same time Transparency of rule structure	C_4 Completeness or Coverage Normalization Distinguishability Complementary

3. C_1: Complexity at the Rule Base Level

In this section, we analyze these criteria that try to reduce or to control the complexity of the RB. The more commonly used measures are the following:

- *Number of rules*: According to the principle of Occam's razor (the best model is the simplest one fitting the system behavior well), the set of fuzzy rules must be as small as possible under conditions in which the model performance is preserved to a satisfactory level.
- *Number of conditions*: The number of conditions in the antecedent of a rule must not exceed the limit of 7 ± 2 distinct conditions, which is the number of conceptual entities a human being can handle [33]. Furthermore, the number of conditions should be as small as possible in order to ease the readability of the rules. Of course, the model performance must also be maintained to a satisfactory level.

Some considerations we should take into account in respect to the aforementioned measures and their descriptions are:

- Differences in simplicity are only remarkable when they are big enough, for example a system with 30 and another one with 32 (or one with 5 and other one with 3) rules, are in practice, at the same level.
- According to the definition of Zhou *et al* in [6], a system must be as simple as possible without seriously affecting its accuracy/usefulness. Nevertheless, even though we agree that having a too simple and very bad system cannot be usefully applied to a real problem, these kinds of systems could allow us to have a general idea of the system's behavior. Generalizations are never good, but a generalization can be useful to show trends.

In the following, we provide a brief review of the approaches that directly take into account the number of rules, conditions, etc, or that include technologies to control or to reduce the complexity at the RB level, for example rule selection.

Ishibuchi *et al* in [16] propose a genetic algorithm for rule selection in classification problems, considering the following two objectives: to maximize the number of correctly classified training patterns and to minimize the number of selected rules. This improves the complexity of the model, thanks to the reduction in the number of rules and the use of "*don't care*" conditions in the antecedent part of the rule. A two-objective genetic algorithm for finding non-dominated solutions in classification problems has also been proposed in [34], with the same two objectives. Then, in [13] they studied both, single-objective and two-objective genetic algorithms, to perform the rule selection on an initial set of classification rules involving "*don't care*" conditions and considering the aforementioned objectives: classification accuracy and number of rules.

An approach to data-based LFM of high-dimensional systems has been proposed by Jin in [35]. He reduces the number of rules, removing redundant rules by means of a fuzzy similarity measure, called similarity of rule premise

(SRP), proposed in [24]. SRP will be explained in depth in the section 5 (quadrant C_3), since it is also devoted to controlling the semantic interpretability at the RB level. This work can also be found in other quadrants, because it uses a distinguishability measure of MFs (C_4) to reduce the number of fuzzy sets (C_2). The characteristics of this work will be explained in their respective quadrants, in section 6 and section 4.

Guillaume in [36] presented a first overview of interpretability-oriented fuzzy inference systems obtained from data and he analyzed different methods from an interpretability point of view. For this quadrant, these methods try to minimize the number of rules by using merging, among others. Moreover, this work analyzed other methods for variable selection and for obtaining the adequate number of MFs, explained in section 4 (quadrant C_2) and it also analyzed a similarity measure as a way of maintaining distinguishability, explained in section 6 (quadrant C_4).

Ishibuchi *et al* in [14] present a Multi-Objective Evolutionary Algorithm (MOEA) for classification problems with three objectives: maximizing the number of correctly classified patterns, minimizing the number of rules and minimizing the number of antecedent conditions. Moreover, they consider two approaches, one for rule selection and a second for rule learning. In [37], they examine the effect of fuzzy partitioning and condition selection in order to find a good trade-off between the number of fuzzy rules and classification performance, by using the said genetic algorithm devoted to maximizing the classification error and to minimizing the number of fuzzy rules. A new approach for regression problems is presented in [38]. They discuss the design of linguistic models with high interpretability considering a fuzzy genetics-based machine learning algorithm and using a Pittsburgh approach. They explain how the formulated linguistic modelling problem can be handled by single-objective and multi-objective genetic algorithms, with three objectives to minimize: total squared error of the rule set, number of fuzzy rules in the rule set and total rule length of the fuzzy rules in the rule set (or total number of antecedent conditions). In the case of the single-objective approach, they used a weighted sum of the three objectives as a fitness function. Moreover, they consider that using "*don't care*" as an additional antecedent fuzzy set is necessary in order to linguistically describe high-dimensional nonlinear functions.

In the following works [39, 40, 41], Cordón *et al* and Alcalá *et al* present models for embedded evolutionary learning of the fuzzy partitions in regression problems. The initial proposal in [39] learns the granularity (number of labels) of the fuzzy partitions and the MFs' parameters (their three parameters jointly). At the same time the authors in [40] propose an evolutionary algorithm to learn the granularity, scaling factors and the domains (i.e., the variable domain or working range to perform the fuzzy partitioning) for each variable. Alcalá *et al* in [41] also propose a method for learning KBs by means of an a priori evolutionary learning of the linguistic fuzzy partition (granularity and translation parameters) that uses the linguistic 2-tuples representation [42]. This methodology allows the reduction of the search space, obtaining more optimal models with high levels of accuracy and simplicity. All these works control the complexity of the RB and linguistic fuzzy partition. In order to do this, they penalize the fitness function with the number of rules obtained, learning models with lower granularities and therefore with a smaller number of rules. The fitness function is defined for minimization:

$$F = w_1 \cdot MSE + w_2 \cdot NR$$

where MSE is the Mean Squared Error (MSE), NR is the number of rules of the obtained KB, $w_1 = 1$ and w_2 is computed from the KB generated from a linguistic fuzzy partition considering the maximum number of labels ($max - lab$, usually 9) and with the MF parameters, $w2 = \alpha \cdot (MSE_{max-lab}/NR_{max-lab})$ with α being a weighting percentage given by the system expert that determines the tradeoff between accuracy and complexity. Values higher than 1.0 search for linguistic models with few rules, and values lower than 1.0 search for linguistic models with high accuracy. Additionally, Cordón *et al* in [43] propose an MOEA for performing feature selection together with linguistic fuzzy partition learning, in order to learn the number of labels for each variable and to adjust the shape of each MF in non-uniform fuzzy partitions, using a non-linear scaling function. They consider the following two objectives: to minimize the classification error percentage and to minimize the complexity (number of features and number of conditions). All these methods seek to decrease the number of MFs. Because of this they are also included in the C_2 quadrant (see section 4).

An overview on the balance between interpretability and accuracy in FM has been presented by Casillas *et al* in [44]. Among others, they analyze different existing methods inside the LFM for reducing the number of linguistic rules and for selecting conditions in the rules. This work also analyzes methods for selecting input variables, for which it has also been included in the C_2 quadrant (see section 4).

Peña-Reyes and Sipper in [33] try to obtain linguistic fuzzy models with a good balance between accuracy and interpretability. To achieve this, they consider several constraints by taking into account both semantic and syntactic criteria, in order to obtain interpretable systems. They propose some strategies to satisfy the semantic and syntactic criteria during the definition of the fuzzy model:

1. Considering linguistic labels, that cover all the variable domain for satisfying the completeness criterion (quadrant C_4).
2. The use of normal, orthogonal MF (quadrant C_4).
3. Allowing "*don't care*" conditions, which reduce the number of antecedents in the rules.
4. Including a default rule, that reduces the number of rules.

This work is also included in quadrant C_4 and it will be explained in section 6.

Guillaume and Charnomordic in [29] present a method for generating interpretable fuzzy rules from data. They include a procedure to simplify an RB in order to get what they call "incomplete rules". These rules are defined only by a few variables and are easier to interpret than "complete rules" (those considering all the system variables). This work is also included in the C_2 quadrant (see section 4), thanks to the reduction in the number of MFs. In [30], they present a fuzzy inference system derivation method together with a simplification algorithm, which includes a mechanism for removing unnecessary rules combined with a procedure for the selection of variables and fuzzy sets. The feature selection and fuzzy set reduction are techniques included in the C_2 quadrant and therefore they will be better explained in section 4. Moreover, this work can also be found in the C_4 quadrant, because it considers the use of a distinguishability measure.

An index comprised of three measures to assess the interpretability of linguistic FRBSs has been proposed by Nauck in [26]. The first measure is the average number of conditions per class in classification problems. The second measure takes into account the coverage of the fuzzy partitions. And the third one measures the number of MFs used in the fuzzy partitions. Because of these last measures, this work is described in depth in quadrant C_2 (section 4) and is also included in quadrant C_4.

Ishibuchi *et al* in [17] apply an improved MOEA, the Multi-Objective Genetic Local Search [32] (MOGLS) for classification problems, considering the same approach as in [14] with three objectives: maximizing the number of correctly classified training patterns, minimizing the number of fuzzy rules, and minimizing the total rule length of fuzzy rules. The approach consists of two phases: first, the method generates candidate rules by using rule evaluation measures and second, the method applies a multi-objective based rule selection. They propose to use two well-known data mining criteria (confidence and support), in order to find a tractable number of candidate fuzzy if-then rules. More specifically, the confidence indicates the degree of the validity of a rule and the support indicates the degree of coverage of a rule.

An automatic method combining different heuristics for designing fuzzy systems from data in classification problems is proposed by Mikut *et al* in [45]. They integrate in the algorithm some components to improve interpretability as follows:

- Generation of rules by decision tree induction and by using a pruning method in order to obtain simple rule conditions and to lead to derived linguistic terms.
- Decreasing the number of generated rules by using an interactive rule selection algorithm, that uses a measure of the relevance, defined by:

$$Q = \underbrace{\left(1 - \frac{E}{E_0}\right)}_{Q_{ac}} Q_{cl}^{\beta} \quad where \quad Q_{cl} = \prod_{r=1}^{r_{max}} max_j(\hat{p}(B_j|P_r)),$$

  where Q is the compromise between classification accuracy (Q_{ac}) and clearness of the rules (Q_{cl}), β is used to control the compromise ($\beta \geq 0$), E is the minimum quadratic error in terms of membership values of the output classes, E_0 is the minimum quadratic error of the trivial model (a rule with an always true premise), r_{max} is the number of rules, and $\hat{p}(B_j|P_r)$ is the probability of "y is in class B_j" for premise P_r.

- Feature selection that allows the reduction of the number of features. A measure to control the complexity at the fuzzy partition level will be explained in section 4 as part of the C_2 quadrant.

A genetic algorithm to perform genetic tuning combined with a fuzzy rule set reduction process that obtains a compact RB with a reduced number of rules has been proposed by Casillas *et al* in [46]. Moreover, they combined linguistic hedges with classic three definition parameter tuning and with domain learning to improve the performance of the system while maintaining its complexity.

Narukawa *et al* in [47] propose an adaptation of the well-known NSGA-II [48] in order to reduce complexity by decreasing the number of rules using three different mechanisms: removing overlapping rules, merging similar rules and by recombining both very different and similar parents. This algorithm includes the following three objectives: maximizing the number of correctly classified training patterns, minimizing the number of fuzzy rules and minimizing the total number of antecedent conditions.

An MOEA for classification problems considering a hybridization of the Michigan and the Pittsburgh approaches is proposed by Ishibuchi and Nojima in [15]. They analyze the interpretability-accuracy tradeoff of fuzzy systems considering three formulations for multi-objective optimization problems (MOPs) and three formulations for single objective optimization problems (SOPs):

- MOP-1: Maximize f_1 and minimize f_2
- MOP-2: Maximize f_1 and minimize f_3
- MOP-3: Maximize f_1, minimize f_2, and minimize f_3
- SOP-1: Maximize $w_1 \cdot f_1 - w_2 \cdot f_2$
- SOP-2: Maximize $w_1 \cdot f_1 - w_3 \cdot f_3$
- SOP-3: Maximize $w_1 \cdot f_1 - w_2 \cdot f_2 - w_3 \cdot f_3$

where w_1, w_2 and w_3 are specified non-negative weights and f_i represents each objective considered: f_1 is the number of correctly classified training patterns, f_2 is the number of fuzzy rules and f_3 is the total number of antecedent conditions of the fuzzy rules, excluding "*don't care*" conditions. The experimental results in [15] demonstrate "the potential advantages of multi-objective formulation over single-objective ones".

Liu *et al* in [31] present a mandani neuro-fuzzy system for balancing interpretability and accuracy. The improvement in interpretability takes place thanks to the reduction in the number of MFs, number of rules and number of attributes. They propose reducing the number of rules by using a method for merging the fuzzy sets and by considering a Hebbian ordering. "The Hebbian ordering is used to represent the importance of the rules, where a higher Hebbian ordering indicates a larger coverage of the training points provided by a given rule. The rules with higher importance are more likely to be preserved" [31]. This work is also included in the C_3 quadrant because they include a mechanism for controlling the consistency of the RB, see section 5.

Alcalá *et al* in [49] propose an effective model of tuning for FRBSs combined with a rule selection, considering the linguistic 2-tuples representation scheme introduced in [42], in order to improve the performance and to decrease the complexity of the classic tuning approaches in complex search spaces. The linguistic 2-tuples allows the lateral displacement of the labels (in fact, the MFs) by considering only one parameter (slight displacements to the left/right of the original MFs). Since the three parameters usually considered per label are reduced to only one symbolic translation parameter, this proposal decreases the learning problem complexity, helping to decrease the model error and facilitating a significant decrease in the model complexity.

An MOEA to obtain a set of solutions with different degrees of accuracy and interpretability has been proposed by Pulkkinen *et al* in [50]. They use the number of misclassifications, the number of rules and the total rule length as objectives to be minimized. Moreover, the authors in [18] present a hybrid genetic fuzzy system [51], which can be used as a reasoning mechanism in a bioaerosol detector. They initialize the population of the MOEA using a decision tree (which implicitly includes features reduction) and include a simplification mechanism in order to reduce the number of rules and the number of rule conditions. They use the following three heuristics in the evolutionary process in order to reduce the complexity:

- Remove all the rules with the same antecedent, except one rule selected randomly.

- The inconsistent rules can be rules of different lengths in which all conditions of the shorter rule(s) are present in the longer rule(s). If these inconsistent rules exist in the RB, they only preserve the longer rule.
- If a condition is present in all the rules, they propose to remove this condition.

In the works [18, 50], the authors apply the C4.5 algorithm [52] to initialize the evolutionary method, which implicitly includes features selection and therefore allows a reduction in the number of variables (quadrant C_2). Moreover, in [18] they also consider mechanisms for merging the fuzzy set (quadrant C_2), for controlling inconsistent rules (quadrant C_3) and for maintaining the distinguishability of the MFs (quadrant C_4), see sections 4, 5 and 6.

Alonso *et al* in [10] propose a methodology for designing highly interpretable linguistic KBs (HILK) for classification problems, considering both expert knowledge and knowledge extracted from data. This methodology includes procedures to merge and to simplify rules, obtaining shorter rules by removing unused variables. In addition, they present a fuzzy interpretability index to quantify the system complexity inspired by the Nauck's index, which combines the following six criteria:

- Total number of rules.
- Total number of premises.
- Number of rules which use one input.
- Number of rules which use two inputs.
- Number of rules which use three or more inputs.
- Total number of labels defined per variable.

These criteria are taken as inputs of a fuzzy system. The fuzzy interpretability index is computed as the result of the inference of a hierarchical fuzzy system made up of four linked KBs generated by HILK. In [8], the authors evaluate the fuzzy interpretability index previously explained. To this end, they perform a very interesting study in which they look for a good interpretability index. In order to do this, they analyze different measures proposed or used by several researchers as the number of rules presented in [16], total rule length originally used in [38], average rule length presented in [38], Nauck's index proposed in [26] and the fuzzy index presented in [10]. This work is also included in the C_3 quadrant, because it uses an analysis of consistency for removing inconsistent rules.

An enhanced MOEA for regression problems has also been proposed in [9] by Alcalá *et al*, and deeply discussed in [12] by Gacto *et al*, which aims to minimize the number of rules together with a classic tuning of the MF parameters (three parameters) by focusing on the most accurate part of the Pareto front in order to find the best trade-off between complexity and accuracy (since both objectives present different levels of difficulty). They used the MSE and the number of rules as objectives to be minimized. Alcalá *et al* in [53] propose an MOEA for learning RBs and parameters of the MFs of the associated linguistic labels concurrently (they use the linguistic 2-tuples representation [42] by only considering one parameter per MF). These works [9, 12, 53] generate a set of FRBSs with different near optimal trade-offs between accuracy and complexity for regression problems.

Pulkkinen and Koivisto in [54] propose "a dynamically constrained multiobjective genetic fuzzy system learning fuzzy partitions, tuning the MFs, and learning the fuzzy rules" for regression problems, considering the following two objectives: MSE and total rule length (sum of the rule lengths). Moreover, the proposed MOEA includes different mechanisms in order to allow a decrease in the number of rules, the number of conditions, the number of MFs, and the number of input variables. This work is also included in the C_4 quadrant because they use constraints to guarantee the distinguishability and the coverage of the fuzzy partitions. See an explanation of these constraints in section 6.

Gacto *et al* in [11] propose a post-processing MOEA to improve the system accuracy while trying to maintain or even improve the interpretability to an acceptable level. This method includes a rule selection mechanism which is combined with a genetic tuning by using the number of rules as an objective to be minimized in order to reduce the model complexity. Because of this (different levels of difficulty in the objectives), it also proposes an enhanced algorithm extending the ideas of the MOEAs in [9, 12]. Moreover, they proposed a semantic interpretability index to be used as a maximization objective (this new semantic interpretability index will be explained in depth in the C_4 quadrant). It was tested on regression problems. Therefore, three objectives are optimized together considering both complexity and semantic interpretability at the same time: MSE minimization, minimization of the number of rules and maximization of the semantic interpretability index.

4. C_2: Complexity at the Level of Fuzzy Partitions

Several measures have been considered in the literature to control the complexity at the level of fuzzy partition. Among them, we should highlight the number of MFs and the number of features. The reduction of these measures also improves the complexity at the RB level. For this reason many of the works in this quadrant can also be found in quadrant C_1, since they are indirectly reducing the number of rules. The most used measures, which are found in this quadrant are:

1. Number of MFs: To control the complexity at the level of fuzzy partitions, it is necessary to have a *moderate number of MFs*. The number of MFs should not exceed the limit of 7 ± 2 distinct MFs, which is the number of conceptual entities a human being can handle [33]. As soon as the number of MFs increases the precision of the system may increase too, but its relevance will decrease.
2. Number of Features or Variables: To reduce the dimensionality in high dimensional problems. The reduction of the number of features can improve the readability of the KB.

In this way, the number of variables as well as the granularity (number of linguistic terms or MFs) of the fuzzy partitions determine the specificity or generality of the model that can be obtained, and they influence proportionally the number of rules of the obtained models. Therefore, C_2 is highly related to C_1. Examples of it are the techniques for rule reduction based on decreasing the number of MFs (merging rules) using similarity measures among MFs.

We have to clarify that, firstly, we will include here these works performing feature selection as a way to reduce the number of features. The selection of conditions has been taken into account as a different technique since it can cause a variable not to be used in one or several rules but it does not disappear from the KB. However, feature selection eliminates the variable from the KB completely. Secondly, we will also include here those works focused on decreasing or controlling the number of MFs of the fuzzy partitions.

Earlier works [27, 28] of Valente de Oliveira propose the use of constraints as a moderate number of MFs in the field of artificial neural networks. To do this, they impose an upper bound on the number of MFs (7 ± 2). In addition, these works propose semantic constraints for the MFs: distinguishability of the MFs, normality and completeness, will be described in the C_4 quadrant (section 6).

The Autonomous Fuzzy Rule Extractor with Linguistic Integrity (AFRELI) algorithm combined with the FuZion algorithm have been proposed by Espinosa *et al* in [22]. The FuZion algorithm allows merging consecutive MFs, in order to reduce the number of fuzzy sets, and to maintain a justifiable number of MFs. This work also maintains the distinguishability between MFs (for more detail see quadrant C_4 in section 6).

Jin in [35] proposes a methodology based on genetic algorithms and the gradient learning method. The author presents a regularization learning to reduce the number of fuzzy sets. "The regularization is to drive the similar fuzzy sets to the same fuzzy set during gradient learning so that the interpretability of the fuzzy system can be greatly improved without seriously deteriorating the system performance". The cost function used for the regularization is defined as:

$$J = E + \gamma\Omega$$

where E is the conventional error function, Ω is the regularization term for merging the similar fuzzy MFs and γ is the regularization parameter ($0 \leq \gamma < 1$). If γ takes high values the system performance is degraded, on the contrary if γ takes low values the interpretability of the fuzzy system is bad. The author assumes that variable x_i has L_i fuzzy subsets $A_i, i = 1, 2, \ldots L_i$. Then they can be divided into m_i (initially equal to L_i) groups using a prescribed similarity threshold δ

$$U_{ik} = \{A_i | S(A_i, A_{k0}) \geq \delta)\}; \quad 1 \leq k \leq L_i$$

where S is the fuzzy similarity measures in [55] (for more details see its definition at the end of this section), U_{ik} denotes a group of fuzzy subsets that are considered to be similar and A_{k0} is the reference fuzzy set for the group. The goal of regularization is to drive the similar fuzzy sets into the same fuzzy set, and the regularization term is defined as:

$$\Omega = \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{m_i} \sum_{A_{ij} \in U_{ik}} (a_{ij} - \bar{a}_{ik})^2 + \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{m_i} \sum_{A_{ij} \in U_{ik}} (b_{ij} - \bar{b}_{ik})^2,$$

$$\text{such that } \bar{a}_{ik} = \frac{1}{I_{ik}} \sum_{A_i \in U_{ik}}^{I_{ik}} a_{ij} \text{ and } \bar{b}_{ik} = \frac{1}{I_{ik}} \sum_{A_i \in U_{ik}}^{I_{ik}} b_{ij}$$

where $\bar{a}_{ik}$ and $\bar{b}_{ik}$ are the parameters of the gaussian function shared by all the fuzzy subsets in group U_{ik}, I_{ik} is the number of fuzzy sets in the group U_{ik} and n is the total number of inputs. The initial values of $\bar{a}_{ik}$ and $\bar{b}_{ik}$ can be the average of the a_{ik} and b_{ik} of subset A_i in the same group U_{ik}. Moreover, this work includes other measures included in the C_1, C_3 and C_4 quadrants, such as number of rules, consistency of the RB and the similarity measure as a way to maintain distinguishability.

An overview of interpretability fuzzy inference systems is presented by Guillaume in [36]. The author analyzes different methods from the specialized literature in order to chose the adequate number of fuzzy sets and to chose the correct methodology for variable/feature selection. This work has also been included in quadrant C_1 and C_4, since it analyzes other measures such as the number of rules and the distinguishability of the MFs.

Casillas *et al* in [44] also analyzed some methods for selecting input variables (feature selection), to obtain the desired balance between the interpretability and accuracy of the fuzzy systems. This work is also included in the C_1 quadrant since they also analyze some methods to improve the complexity at the RB level.

In the following works [39, 40, 41, 43], Cordón *et al* and Alcalá *et al* present different methodologies to obtain the whole KB of an FRBS based on the embedded genetic learning of linguistic fuzzy partitions, considering an evolutionary process that learns the number of MFs. In section 3 (quadrant C_1), a deeper explanation of these works is included, which makes use of a penalization by the number of rules in order to control the complexity (learning fuzzy partitions with low granularities).

Tikk *et al* in [56] present an algorithm for feature selection in order to reduce the complexity in classification problems. They search for features which maximize the average distance between the classes. They propose the use of the Sequential Backward Selection [57] as a search method in order to rank the features. This method removes one feature in each stage of the search process, in this way reducing complexity.

A feature selection algorithm which should be useful in height dimensional classification problems is proposed by Vanhoucke *et al* in [58]. This algorithm ranks the input features according to their mutual information, and discards all features deemed irrelevant by a threshold criterion.

Nauck in [26] presents an index for classification problems to measure the interpretability (I) of linguistic fuzzy models, in terms of complexity ($comp$), the degree of coverage ($\overline{cov}$) of the fuzzy partition, and a partition complexity measure ($\overline{part}$) that penalizes partitions with a high granularity. Thanks to this last measure ($\overline{part}$) it tries to obtain a fuzzy rule based system with a small number of fuzzy sets. This index is defined as:

$$I = comp \cdot \overline{cov} \cdot \overline{part}.$$

In the following, these measures are formulated. The complexity measure ($comp$) is defined as:

$$comp = m \Big/ \sum_{i=1}^{r} n_i$$

where m is the number of classes, r is the number of rules and n_i is the number of variables used in the i-th rule (also included in C_1 quadrant).

The degree of coverage $\overline{cov}$ is defined as the average normalized coverage on cov_i:

$$cov_i = \frac{\int_{X_i} \bar{h}_i(x)dx}{N_i} \quad where \quad \bar{h}_i(x) = \begin{cases} h_i(x) = \sum_{k=1}^{p_i} \mu_i^{(k)}(x), & \text{if } 0 \le h_i(x) \le 1 \\ \frac{p_i - h_i(x)}{p_i - 1}, & \text{otherwise} \end{cases}$$

where X_i is the domain of the i-th variable and this domain is partitioned by p_i fuzzy sets and with $N_i = \int_{X_i} dx$ for continuous domains. This measure tries to maintain the semantic interpretability at the fuzzy partition level, and therefore it will also be mentioned as a work related to quadrant C_4.

The partition complexity measure $\overline{part}$ is the average normalized partition measure on $part_i$:

$$part_i = 1 \Big/ p_i - 1$$

where p_i is the number of fuzzy sets in the i-th variable.

Guillaume *et al* in [29, 30] present a sophisticated distance function, with external and internal distances, which is used to merge fuzzy sets, thus allowing a moderate number of MFs for classification problems to be obtained. They propose the application of the merging of fuzzy sets that minimizes the variation of the D_m index defined for a given size m partition as:

$$D_m = \frac{1}{N(N-1)} \sum_{q,r=1,2,\ldots N, q\neq r} d(q,r)$$

where m is the number of fuzzy sets in the fuzzy partition, N is the number of training data and the pairwise distance $d(q,r)$ will take into account the memberships of the different q and r training points by combining the respective parts of internal and external distances. Internal and external distances as well as the pairwise distance $d(q,r)$ are defined in the following.
The *internal distance* is a measure of membership similarities for a given fuzzy set f and it is computed, given two data points with (x_q^j, x_r^j) coordinates for the j-th dimension, by means of the difference of the membership degrees:

$$d_{int}^f(q,r) = |\mu_q^f - \mu_r^f|$$

The *external distance* is a measure combining internal distances and prototype distances. It takes into account the point location within the fuzzy set and the relative position of the fuzzy set in the fuzzy partition. The external distance between two points which belong to the f and g fuzzy sets respectively is defined as:

$$d_{ext}^{f,g}(q,r) = |\mu_q^f - \mu_r^g| + d_{prot}(f,g) + D_c$$

where D_c is a constant correction factor, which ensures that the external distance is always superior to any internal distance, and $d_{prot}(f,g)$ is the prototype distance (numerical or symbolic distance) between the centers of fuzzy sets f and g.
Taking into account that $d_{f,g}(q,r)$ represents respective memberships of the fuzzy sets f and g, and that it is an internal distance if $f = g$ or an external distance otherwise, $d(q,r)$ is defined as:

$$d(q,r) = \frac{1}{\sum_{f=1}^{m} \mu_q^f} \sum_{f=1}^{m} \left[\mu_q^f \frac{1}{\sum_{g=1}^{m} \mu_r^g} \sum_{g=1}^{m} [\mu_r^g \, d_{f,g}(q,r)] \right]$$

These works [29, 30] include a mechanism for removing unnecessary rules, therefore they are also included in the C_1 quadrant. Moreover, the proposal in [30] will be also mentioned as a work related to quadrant C_4, since it also considers a distinguishability measure.

In [45], Mikut *et al* proposed an automatic method for designing fuzzy systems from data, by using a decision tree, rule selection and feature selection. Feature selection is used to reduce the number of features by determining the most important features. To do this, they present a feature relevance measure that reflects the preference and relevance of a feature:

$$M_l = M_{l,ap}^{\alpha} \frac{H(x_l; y)}{H(y)} \;\; where \;\; H(y) = -\sum_{j=1}^{m_y} p(B_j)\, lg(p(B_j)),$$

where $H(y)$ is the entropy of the output y, $H(x_l; y)$ is a measure of the average information provided by feature x_l about the class of y, $M_{l,ap}$ is a relevance weight provided by the user, α is the strength of the feature preference (if α is near to zero the influence of a priori preferences diminishes), m_y is the number of classes, $p(B_j)$ is the probability of the event "y is in class B_j" and and lg is the logarithm in base 2. This work is also included in quadrant C_1 because of the rule selection technique.

Pulkkinen *et al* in [18, 50] first apply the C4.5 algorithm [52] in order to create a decision tree and then it uses the decision tree to obtain a fuzzy classifier. The C4.5 algorithm implicitly includes a mechanism for reducing the number of features. Both works are also mentioned in the C_1 quadrant, because they include mechanisms to reduce the number of rules and the number of rule conditions. Moreover, the authors present in [18] an MOEA that uses the similarity measure (S), proposed in [55], for merging fuzzy sets. This measure is a fuzzy relation that expresses the degree to which A and B are equal and is defined as follows:

$$S(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|},$$

where intersection ($\cap$) and union ($\cup$) are defined by a proper couple of t-norm and t-conorm and $|\cdot|$ is the cardinality of the resulting fuzzy set. If the similarity measure is greater than a given threshold (a suitable value for the threshold is 0.25), then they merge these two fuzzy sets (A and B) to generate a new one C. The merging method creates a common trapezoidal fuzzy set C that replaces the occurrence of the merged trapezoidal fuzzy sets A and B, defined as $\mu_A = (x; a_1, a_2, a_3, a_4)$ and $\mu_B = (x; b_1, b_2, b_3, b_4)$. The fuzzy set C is defined as $\mu_C = (x; c_1, c_2, c_3, c_4)$ where:

$$c_1 = min(a_1, b_1); \;\; c_2 = \lambda_2 a_2 + (1 - \lambda_2) b_2; \;\; c_3 = \lambda_3 a_3 + (1 - \lambda_3) b_3; \;\; c_4 = max(a_4, b_4),$$

and the parameters $\lambda_2, \lambda_3 \in [0,1]$ determine which of the fuzzy sets A or B has the highest influence on the kernel of C. This work will be also mentioned as a work related to quadrants C_3 and C_4, because it also controls the consistency of the RB and its distinguishability.

5. C_3: Semantic at the Rule Base Level

Assuming that the initial fuzzy partitions are interpretable at the semantic fuzzy partition level, this quadrant is related to measures or properties that are devoted to controlling the semantic interpretability at RB level. This is the quadrant in which there are fewer works in the literature. Mainly, this quadrant takes into account the following properties:

- *Consistency of the RB*, which means the absence of contradictory rules in RB, in the sense that rules with similar premise parts should have similar consequent parts.
- *Number of rules fired at the same time*, which consists of minimizing the number of rules firing that are activated for a given input.

Most of the works in this quadrant are focused on consistency. By contrast, there are only a few works that make use of the number of rules fired, among them the works by Chen *et al* in [21, 59]. However, in our opinion this measure represents a very promising way to preserve the individual meanings of the linguistic rules comprising the RB. In the following, we briefly describe those works that look to improve the interpretability in this quadrant.

Jin *et al* in [24, 60] propose a methodology for generating flexible, complete, consistent and compact fuzzy rule systems from data using evolutionary algorithms. They propose some indices for coverage and consistency of the linguistic fuzzy system and they integrate them into an objective function by means of an aggregated function f. The consistency index ($Cons$) is calculated as follows for two given rules $R(i)$ and $R(k)$:

$$Cons(R(i), R(k)) = exp\left\{ - \frac{\left(\frac{SRP(i,k)}{SRC(i,k)} - 1.0\right)^2}{\left(\frac{1}{SRP(i,k)}\right)^2} \right\}$$

where SRP is the similarity of rule premises and SRC is the similarity of rule consequents, and they are calculated as:

$$SRP(i,k) = \min_{j=1}^{n} S(A_{ij}, A_{kj}), \quad SRC(i,k) = S(B_i, B_k)$$

where S is the similarity measure between two fuzzy sets (A and B) proposed in [55], which was described in the previous section. Finally, the degree of inconsistency of a given RB (f_{Incons}) is calculated as:

$$f_{Incons} = \sum_{i=1}^{N} Incons(i) \,,$$

where $Incons(i)$ is the degree of inconsistency for the i-th rule. It is defined as:

$$Incons(i) = \sum_{\substack{1 \leq k \leq N \\ k \neq i}} [1.0 - Cons(R^1(i), R^1(k))] \; +$$

$$+ \sum_{\substack{1 \leq l \leq L \\ i = 1,2,\ldots,N}} [1.0 - Cons(R^1(i), R^2(l))],$$

where R^1 and R^2 denote the RB generated from data and the RB extracted from prior knowledge (since the authors defined this index considering the possibility of including rules provided by experts) and N and R are the rule number of R^1 and R^2, respectively.

The proposed approach is applied to the design of a distance controller for cars. In this problem the objective function f is formulated as:

$$f = f_E + \xi \cdot f_{Incons} + f_{Incompl} \;\; where \;\; f_E = \sum_{t=1}^{J} \sqrt{(v(t) - v_d(t)^2)},$$

J is the total number of sampled data, $v_d(t)$ is the target velocity, $v(t)$ is the velocity of the controlled car, ξ is a weighting constant to control the consistency level and $f_{Incompl}$ is a penalization constraint to maintain the completeness of the fuzzy partition, for more details see the C_4 quadrant in section 6.

In [35], Jin makes use of the previously explained measure (SRP) to control the consistency of the RB. Moreover, this work includes other measures explained in the quadrants C_1, C_2 and C_4, such as number of rules, the number of fuzzy sets and a distinguishability measure.

Cheong and Lai in [21, 59] present a parallel genetic algorithm for obtaining a fuzzy logic controller with some constraints in the RB. This algorithm tries to minimize the number of rules fired at the same time, which is not only devoted to improving the consistency of the RB, but also to reducing the effort needed to understand the meaning of the rules since they present less interactions among them. This helps to provide rules that can be better understood by human beings. Moreover, the authors use semantic constraints in order to guarantee the distinguishability of the MFs. Therefore, these works are also included in the C_4 quadrant.

Pedrycz in [61] analyzes the interpretability of an RB by using two measures, relevance and consistency:

1. The relevance of a rule "is quantified in terms of the data being covered by the antecedent and conclusions standing in the rule". This measure is defined by the author as:

$$rel(A_i \times B_i) = \sum_{k=1}^{N} A_i(x_k) \; t \; B_j(y_k),$$

   where A_i is the antecedent of the i-th rule, B_j is the consequent of the rule, N is the number of data and t is the t-norm used. This measure presents higher values when the rule has more relevance in the RB.
2. The consistency of the rule "expresses how much the rule interacts with the others in the sense that its conclusion is distorted by the conclusion parts coming from other rules". This measure is defined as:

$$Cons(i, R) = \sum_{j=1, j\neq i} cons(i, j),$$

$$cons(i, j) = \frac{1}{N} \sum_{k=1}^{N} A_i(x_k) \equiv A_j(x_k) \rightarrow (B_i(y_k) \equiv B_j(y_k)),$$

   where i and j are the indexes of the rules in R and $\rightarrow$ is the implication operator used.

A neuro-fuzzy method, more deeply explained in quadrant C_1, is proposed by Liu *et al* in [31]. This methodology reduces the number of rules by merging MFs. This reduction can provoke the appearance of inconsistent rules. However, the authors solve the problem of inconsistent rules by maintaining only the most important rules when this conflict arises.

Pulkkinen *et al* in [18] include a mechanism (previously explained in the C_1 quadrant) to prevent the RB from having inconsistent rules. This work is also mentioned as a work related to quadrants C_1, C_2 and C_4, because it also controls the number of rules, the total rule length, the consistency of the RB and the distinguishability of the MFs.

Alonso *et al* in [10] propose the HILK methodology, which also takes into account the conflicting rules, using the following solutions:

- Totally inconsistent rules are rules with the same premises and different conclusions. These rules will be removed from the RB.
- Specialization rule is a rule in which the space of the premises are included in another rule, and both rules have different conclusions, i.e., a specific rule is a specialization of the most general one. There are two different ways to avoid this contradiction:
  - Keep only one rule corresponding to the largest input domain and set the best suited consequent.
  - Keep the most specific rule and split the most general one, to cover the same input domain except the one covered by the specific rule.
- Partially inconsistent rules are rules with no empty intersection in the premises and with different consequents. The way to avoid the intersection in the premises is splitting these rules and choosing the best suited consequent.

This work is also included in the C_1 quadrant because it also considers the number of rules and the number of conditions.

Alonso *et al* in [8] propose a methodology in order to analyze different measures by using a web poll dedicated to determining how different people assess interpretability by giving priority to different criteria. The authors consider ten variables as tentative interpretability indicators: number of rules, total rule length, number of inputs, number of labels used in the RB, percentage of rules which use less than ten percent of inputs, percentage of rules which use between ten and thirty percent of inputs, percentage of rules which use more than thirty percent of inputs, percentage of elementary labels used in the RB, percentage of OR composite labels used in the RB and percentage of NOT composite labels used in the RB. Finally, they conclude that the results extracted from the poll show the inherent subjectivity of the measures, obtaining a huge diversity of completely different answers. However, "three interpretability indicators turn up as the most significant ones: total rule length, number of used labels in the RB, and percentage of NOT composite linguistic terms". In fact, the use of NOT composite linguistic terms or even OR disjunctive operators could also affect the transparency of the rule structure.

In this quadrant, there are few measures to quantify the semantic interpretability at level of RB. It would be interesting to propose new measures for this C_3 quadrant or to fix the most appropriate from the existing ones. As mentioned, some additional aspects could be to measure "the percentage of OR composite labels used in the RB" and/or "the percentage of NOT composite labels used in the RB" as Alonso *et al* in [8] suggest from the studies based on the web pool.

6. C_4: Semantic at the Fuzzy Partition Level

In this quadrant, we look to maintain semantic interpretability at the fuzzy partition level. From a classical point of view, this problem has been tackled by applying some constraints to the MF's definition in order to preserve or to improve some desirable properties. Some of the most important properties defined by the experts in this framework are:

1. *Completeness or Coverage*: The universe of discourse of a variable should be covered by the MFs, and every data point should belong to at least one of the fuzzy sets and have a linguistic representation, i.e., it is required that membership value should not be zero for all the linguistic variable domains.

2. *Normalization*: MFs are normal if there is at least one data point in the universe of discourse with a membership value equal to one, in respect to the maximum membership degree.
3. *Distinguishability*: An MF should represent a linguistic term with a clear semantic meaning and should be easily distinguishable from the remaining MFs of the corresponding variable.
4. *Complementarity*: For each element of the universe of discourse, the sum of all its membership values should be near to one. This guarantees a uniform distribution of the meanings among the elements.

Taking into account these properties and the semantic constraints classically used to preserve them, we can observe that they represent absolute properties that try to obtain well distributed MFs, ensuring complementary membership degrees between each two adjacent concepts. In fact, *strong fuzzy partitions* satisfy these semantic properties (distinguishability, coverage, normality, complementarity, etc) to the highest level. These kinds of fuzzy partitions, in which the sum of the membership degrees within the variable domain are equal to 1.0 and the MFs are equidistant (therefore, also symmetrical), perfectly meet the required semantic constraints and they are widely assumed to have high semantic interpretability. However, it is not always possible to use strong fuzzy partitions or to impose absolute properties. If the system is provided from expert knowledge, the experts can consider another type of fuzzy partitioning more appropriate for the problem. Therefore, it would be interesting to take into account relative measures which try to measure the interpretability with respect to the fuzzy partitions intended as the most interpretable ones.

As mentioned above, some of the works that are in this quadrant look to maintain interpretability, introducing *semantic constraints* in the modelling process. By contrast, some others are devoted to proposing semantic interpretability *measures* to quantify or to optimize the semantic interpretability properties. Those works that consider constraints or even measures but simply for setting limits on some values of the MFs are included in subsection 6.1. Those works that propose measures to quantify the fuzzy partition interpretability are studied in subsection 6.2.

6.1. Semantic Interpretability Constraints at the Fuzzy Partition Level

This subsection presents those works that use measures to impose constraints on the MFs or that directly control the limits on the values of the MFs at the fuzzy partition level. In the following, we shortly review these approaches.

The coverage and the distinguishability of the fuzzy partitioning of each input variable is examined using the fuzzy similarity measure (S) [55] (previously explained in section 4) by Jin *et al* in [24, 60]. This value ranges from 0 to 1. To keep the MFs with a proper shape, the fuzzy similarity measure of any two neighboring MFs is required to satisfy the following constraint:

$$FSM^{-} \leq S(A_i, A_{i+1}) \leq FSM^{+}$$

where A_i and A_{i+1} are two neighboring fuzzy sets and both FSM^- and FSM^+, are the desired lower and upper bound of the fuzzy similarity measure respectively. It will be also mentioned as a work related to quadrant C_3, because it proposes measures to control the consistency of the RB. Additionally, Jin in [35] proposes a measure for finding similar fuzzy sets. This measure is based on a distance measure, $d(A_1, A_2)$ based on the existence of gaussian MFs (with a_i and b_i the core and width definition parameters for the MF A_i):

$$S(A, B) = \frac{1}{1 + d(A_1, A_2)}, \;\; where \;\; d(A_1, A_2) = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2}$$

Moreover, this work includes other measures explained in the C_1, C_2 and C_3 quadrants, such as number of rules, the number of fuzzy sets and the consistency measure.

In [22], Espinosa and Vandewalle propose an algorithm, named FuZion, to merge MFs whose cores are too close to each other. The proposed algorithm guarantees the distinguishability and the coverage properties by imposing some constraints. A fundamental parameter of this algorithm is the minimum acceptable distance (M) between modal values. When the value of M is smaller, the number of acceptable MFs per domain will increase, increasing the number of rules and also increasing the complexity of the model. On the other hand, as the value of M increases the number of MFs per domain decreases, reducing the number of rules and increasing the approximation error. This parameter that must be used to balance the tradeoff between interpretability and precision, should be fixed to values between 5 – 25% of the coverage of the universe of discourse to guarantee semantic integrity. The FuZion algorithm maintains a justifiable number of MFs, for a more detailed description see the quadrant C_2 in section 4.

Cheong *et al* in [21, 59] present several semantic constraints that are applied to the optimization process in order to produce well-formed fuzzy sets. In this parallel genetic algorithm, they use strong fuzzy partitions with triangular MFs and only one parameter is used for evolutionary adaptation of the MFs. This is the location of the center of the triangle while the remaining parameters are moved automatically since the last point of each MF is kept equal to the center point of the following MF. In a normalized universe of discourse in [-1, 1], the first and last triangles were fixed to -1 and 1, respectively, and the other triangles constrained to the range $[X_1^s, X_2^s]$, which is defined as follows for the center of the s-th triangle:

$$[X_1^s, X_2^s] = (2/n * (s-1) \pm a)$$

where n is the number of fuzzy sets in the universe of discourse, $1 < s < n$, and a is constant with a value determined experimentally (an adequate portion of the discourse universe). These semantic constraints are used in order to guarantee distinguishability. These two works are also mentioned in the C_3 quadrant, since they propose measures to control the number of rules firing at the same time and the consistency of the RB.

Mencar *et al* in [25] have focused on defining a faster alternative to the similarity metrics as a way to measure the distinguishability of the MFs. The most common measure to quantify distiguishability is the similarity S [55], previously explained in section 4. The problem of the similarity measure S is that its calculation is usually computationally intensive. For this reason the authors propose in [25] the use of a possibility measure (Π) as an alternative whose calculation can be very efficient. This measure between two fuzzy sets A and B is defined as follows:

$$\Pi(A, B) = \sup_{x \in U} min\{\mu_A(x), \mu_B(x)\}$$

The possibility measure can be also used to evaluate distinguishability. As for similarity measure $S(A, B)$, distinguishability can be formally defined as the complement of the possibility between two distinct fuzzy sets, which must not be less than a predetermined threshold δ:

$\Delta(A, B) = 1 - \Pi(A, B) \geq \delta$ which implies $\forall A, B \in F : A \neq B \longrightarrow \Pi(A, B) \leq \vartheta$ with $\vartheta = 1 - \delta$

Pulkkinen and Koivisto in [54] present an MOEA for solving regression problems with two objectives: MSE and number of conditions (for which it can also be found in the C_1 quadrant). Since they are also evolving MF parameters, they propose dynamic constraints in order to guarantee the transparency (distinguishability and coverage) of fuzzy partitions, using the following conditions:

- Symmetry conditions: The symmetry is guaranteed by definition since they use Gaussian MFs.
- α-condition: To control the intersection point of two MFs. At any intersection point of two MFs, the membership values is at most α.
- γ-condition: To control the overlapping in the center of each MF. At the center of each MF, no other MF receives membership values larger than γ.
- β-condition: To ensure that the Universe of Discourse is strongly covered. At least one MF has its membership value at β.

These constraints must be fixed previously in order to apply the dynamic tuning strategies. The authors recommend as appropriate values the following ones: $\alpha = 0.8$, $\gamma = 0.25$ and $\beta = 0.05$.

6.2. Semantic Interpretability Measures at the Fuzzy Partition Level

This subsection presents those works that propose or use a measure that allows the semantic interpretability of the fuzzy partitions obtained by the different learning techniques used with this aim to be quantified. In the following they are briefly described.

Valente de Oliveira *et al* in [27, 28, 62] propose some semantic constraints for MFs together with some interpretability metrics or measures, including distinguishability of MFs, moderate number of MFs, natural zero positioning, normality and completeness. Almost all the most used and accepted constraints, or absolute characteristics were proposed in these first works. In this way, the authors try to avoid potential inconsistencies in linguistic fuzzy models. They propose the following expressions for coverage(J_1), and for distinguishability (J_2), and they use them to enforce the interpretability of fuzzy systems during the gaussian MFs' optimization problem using a backpropagation algorithm. These measures are defined for a given external variable as:

$$J_1 = \frac{1}{2}\sum_{k=1}^{N}(x[k] - \bar{x}[k])^2 \;\; where \;\; \bar{x}[k] = \frac{\sum_{i=1}^{n} \mu_i(x[k])a_i}{\sum_{i=1}^{n} \mu_i(x[k])},$$

$$J_2 = \frac{1}{2}\sum_{k=1}^{N}\left[\left(\sum_{k=1}^{N} \mu_i^p(x[k])^{i/p} - 1\right)\right]^2,$$

N is the number of training data, n is the total number of elements MFs, a_i $(i : 1 \ldots n)$ are the centers of the generic MFs μ_i, $x[k]$ is the k-th numeric sample and p is used to control the strength of the J_2 measure. If $p = 1$ they have a strong influence whereas it can be eliminated as $p \longrightarrow \infty$.

In [27, 28, 62], they use a linear combination of the two constraints as a fitness function:

$$\overline{J} = J + K_1 J_1 + K_2 J_2$$

where K_1 and K_2 are positive penalty factors and J is the MSE. Moreover, these works [27, 28] impose an upper limit on the number of MFs. For a more detailed description see quadrant C_1 in section 3.

An aggregation procedure that merges fuzzy sets using similarity measures is proposed by Guillaume *et al* in [30]. The merging procedure as well as the similarity index used have been described in section 4, quadrant C_2. Additionally, to assess the validity of a fuzzy partition they propose a new index based on the homogeneity of the fuzzy set densities. The density d_f for a fuzzy set f is equal to the ratio of its weight, or fuzzy cardinality w^f defined as $w^f = \sum_{x \in E} \mu_j^f(x)$, divided by the fuzzy set area, where E is the subset of learning samples covered by f. The density homogeneity σ^{FP}, is defined as the density standard deviation for all the fuzzy sets of the fuzzy partition:

$$\sigma^{FP} = \sqrt{(1/m)\sum_{f=1}^{m}(d_f - \bar{d})^2},$$

where $\bar{d}$ is the mean of the fuzzy set densities. From the homogeneity point of view the best partition is the one for which σ^{FP} reaches a minimum. This work has also been included in quadrants C_1 and C_2, because it uses mechanisms for removing unnecessary rules, for feature selection and for reducing the number of MFs. Guillaume in [36] also analyzes the similarity measures as a way of maintaining the distinguishability of the fuzzy set within this overview work. Moreover, since this is an overview work it can also be found in the quadrants C_1 and C_2, considering the use of merging methods and feature selection methods (see sections 3 and 4 for more details).

Furuhashi *et al* and Susuki *et al* in [63, 64, 65], propose a conciseness measure based on the combination of De Luca and Termini's fuzzy entropy [66] and a measure for the deviation of an MF. They consider that "a fuzzy model is more concise if the MFs are more equidistantly allocated in the universe of discourse, and the shapes of MFs are less fuzzy". De Luca and Termini's fuzzy entropy, $d(A)$, can be used to evaluate the shapes of the MFs. The fuzzy entropy of a fuzzy set A is defined as:

$$d(A) = \int_{x_1}^{x_2} \{-\mu_A(x)\; ln(\mu_A(x)) - (1 - \mu_A(x))\; ln(1 - \mu_A(x))\} dx,$$

where x_1 and x_2 are the left and right points of the support of the fuzzy set A, and $\mu_A(x)$ is the MF of the fuzzy set A. If $\mu_A(x) = \frac{1}{2}$ for all x on the support of A, then the fuzzy entropy of the fuzzy set A is the maximum. On the other hand, the authors define the measure for the deviation of an MF, $r(A)$, as the quantitative measure of the deviation of an MF from the symmetry and it is given as:

$$r(A) = \int_{x_1}^{x_2} \mu_C(x)\; ln\left(\frac{\mu_C(x)}{\mu_A(x)}\right) dx,$$

where $\mu_C(x)$ is the symmetrical MF associated to the fuzzy set A which has the same support than the fuzzy set A. Finally, they define the average conciseness measure as:

$$dr_{avr} = \frac{1}{N_m - 2}\sum_{i=2}^{N_m - 1} dr(A_i), \;\; where \;\; dr(A) = d(A) + r(A) = -\int_{x_1}^{x_2} \mu_C(x)\; ln(\mu_A(x)) dx,$$

and N_m is the number of fuzzy sets in the universe of discourse. They use an MOEA considering the following two objectives: the accuracy of the model and the average conciseness, which is previously explained.

Nauck in [26] presents an index I with measures for controlling the complexity ($comp$) of the RB, (also included in the C_1 quadrant), the coverage ($\overline{cov}$) of the fuzzy partition and the number of MFs in the fuzzy partitions ($\overline{part}$) for classification problems. Usually, a fuzzy partition is considered to be "good", if it provides complete coverage (i.e. membership degrees add up to 1 for each element of the domain). In the C_2 quadrant, a deeper explanation of this work is included (see section 4).

Fazendeiro *et al* in [67] propose an interpretability measure (J), for multi-objective and mono-objective optimization, that is obtained by means of the aggregation of three indices:

- The first index J_1 is intended to promote the natural localization of the linguistic term Zero.
- Index J_2 penalizes MFs with a poor distinguishability level.
- Index J_3 penalizes the low level of coverage of the universe of discourse.

Additionally, they use normal MFs to satisfy the normalization property. The three indices and the aggregated measures are defined as:

$$J_1 = K_1\, c_{ZE}^2 \; step\; (c_{ZE}),$$
$$J_2 = K_2 \times \sum_x \left[(M_p(A)-1)^2 step(M_p(A)-1)\right],$$
$$J_3 = K_3 \times \sum_x \left[(M_p(A)-1)^2 step(M_p(A)-1)\right],$$
$$J = \sum_i J_i\ ,$$

where c_{ZE} denotes the center of the membership function of the linguistic term $Zero$, K_1, K_2 and K_3 are constants which allow the tuning of the relative weight of each J_i, the function $step$ is the standard unit step function and the sigma-count operator $M_p(A)$ gives a measure of the cardinality of a fuzzy set A, as follows:

$$M_p(A) = \sqrt[p]{a_1^p + \ldots + a_n^p},$$

where a_i ($i = 1, \ldots, n$) are the degrees of membership defining a fuzzy set A and p is a positive integer (in the experiments they used $p = 1$). In other works [68, 23] of the same authors, they also use the indices previously presented to guarantee interpretability, focusing on solving a control of neuromuscular blockade problem. Moreover in [23], they consider the use of an MOEA to optimize error and J separately.

Several metrics to guarantee the semantic interpretability of the MFs are presented by Pulkkinen *et al* in [18]. The proposed semantic interpretability metrics are:

1. Overlap penalty (P_{OP}): It is the length of overlap for fuzzy sets,

$$P_{OP} = \frac{1}{n_s}\sum_{i=1}^{n_s}\frac{1}{N_{ov}^i}\sum_{j=1}^{N_{ov}^i}\frac{\lambda_{i,j}}{\chi_i} \;\; where \;\; N_{ov}^i = \binom{M_i}{2}\frac{M_i!}{2(M_i-2)!},$$

   where n_s is the number of variables selected from the n variables in the initialization, $\lambda_{i,j}$ is the length of the j-th overlapping between two MFs for the input variable i, N_{ov}^i is the number of MF pairs in the input variable i, $M_i \geq 2$ is the number of active fuzzy sets in the input variable i (if there are only 2 MFs, $N_{ov}^i = 1$) and $\chi_i = ubound_i - lbound_i$ ($ubound_i$ and $lbound_i$ are, respectively, the upper and lower bounds of the i-th variable). P_{OL} is not calculated for a certain variable, if the number of active MFs assigned to it is less than 2.

2. Discontinuity penalty (P_{DC}): It is the proportion of total length of discontinuity for two fuzzy sets.

$$P_{DC} = \frac{1}{n_s}\sum_{i=1}^{n_s}\sum_{j=1}^{G_i}\frac{\psi_{i,j}}{\chi_i}$$

   where G_i is the number of discontinuities and $\psi_{i,j}$ is the length of the j-th discontinuity in the input variable i.

3. Middle value penalty (P_{MV}): It is used to prevent relaxed covering of MFs.

$$P_{MV} = \frac{1}{n_s} \sum_{i=1}^{n_s} \delta_i \ \ where \ \ \delta_i = \begin{cases} \frac{\delta_i^* - \alpha_L}{1-\alpha_L} & \text{if } \delta_i^* > \alpha_L \\ 0 & \text{if } \delta_i^* \leq \alpha_L \end{cases}$$

where α_L is the user specified maximum covering level ($\alpha_L = 0.1$) allowed for an MF in the center definition point of another MF, and δ_i^* is the maximum covering of other MFs at any center definition point of the MFs in variable i. P_{MV} is not calculated for a certain variable, if the number of active MFs assigned to it is less than 2.

They define a transparency penalty (T) which is used to control distinguishability by including this measure T as an objective in the proposed MOEA together with the misclassification rate (as a metric of accuracy, true positive and false positive rates). T is defined as follows:

$$T = P_{OL} + P_{DC} + P_{MV}.$$

This work also considers measures to minimize the number of rules and the number of premises (quadrant C_1) and it also includes mechanisms for merging the fuzzy set and for feature selection (quadrant C_2), for controlling inconsistent rules (quadrant C_3) (see sections 3, 4 and 5 for more details).

An MOEA to perform context adaptation is presented by Botta *et al* in [20]. This algorithm considers the system error (MSE in regression problems) and an interpretability index ($\Phi_Q(P)$) to preserve the fuzzy ordering and good distinguishability by using scaling functions: core-position modifier, core-with modifier, support-width modifier and generalized positively modifier (i.e., changing the degree of membership of the boundary elements of the fuzzy sets). This interpretability index considers a fuzzy partition, $P = A_1, \ldots, A_i, \ldots, A_N$, consisting of N fuzzy sets where $d_{j,i} = |j - i|$ is the semantic distance between A_j and A_i. For instance, the semantic distance between A_3 and A_1 is 2. The index is defined as:

$$\Phi_Q(P) = \frac{\sum_{1 \leq i \leq N-1} \sum_{i<j \leq N} \frac{1}{d_{j,i}} \cdot \mu_Q^{d_{j,i}}(\overbrace{Q \leq (A_i, A_j)}^{x})}{\sum_{1 \leq i \leq N-1} \sum_{i<j \leq N} \frac{1}{d_{j,i}}}$$

where Q is a fuzzy ordering index, $x = Q \leq (A_i, A_j)$ and $\mu_Q^{d_{j,i}}(x)$ are fuzzy sets of the values of Q defined in the universe [0,1].

Gacto *et al* in [11] propose an index (namely $GM3M$) that helps preserve the semantic interpretability of linguistic fuzzy systems. This index is devoted to maintainining the original shape of the MFs while a tuning (or any kind of learning or improvement) of their definition parameters is performed, and it represents a relative measure of the quality of the linguistic fuzzy partitions, once we know how the most interpretable ones should be.

Therefore a possible way to measure the interpretability of the MFs is measuring it with regard to the strong fuzzy partitions (which usually satisfy absolute semantic constraints or absolute measures to the highest degree). On the other hand, since the concepts and their meaning strongly depend on the problem and person who makes the assessment (the final user), the initial linguistic fuzzy partitions could also be given by an expert. The index ($GM3M$) allows for work with strong fuzzy partitions or with linguistic fuzzy partitions defined by an expert, being a relative index capable of quantifying the interpretability of the fuzzy partitions with respect to the original ones, solving the problem of absolute measures when the expert has the concepts clear and they do not fit with the imposed absolute properties.

$GM3M$ is defined as the geometric mean of three metrics, and its values range between 0 (the lowest level of interpretability) and 1 (the highest level of interpretability). The index is defined as:

$$GM3M = \sqrt[3]{\delta \cdot \gamma \cdot \rho}$$

where δ, γ and ρ are three complementary metrics to measure interpretability when a tuning is performed on the MFs, i.e., when the MF definitions are changed, which is usually needed to reach an acceptable accuracy level. The different metrics allow the authors to ensure the interpretability to a minimum level in all the MFs, since their main aim is to measure the worst case. Therefore, if there is an important problem in any of the MFs, it can be detected and reflected in each particular metric. The geometric mean is used in case only one of the metrics has very low values (causing low interpretability) small values of $GM3M$ are also obtained. Each metric was proposed for work with triangular MFs but they can easily be extended with some small changes in the formulation to gaussian or trapezoidal MFs (see [11] for more details on these extensions). In this way, it represents the first almost general index to quantify the semantic interpretability of fuzzy partitions. The said metrics are:

- MFs displacement (δ): This metric measures the proximity of the central points of the MFs to the original ones.
- MFs lateral amplitude rate (γ): This metric measures the left/right rate differences of the tuned and the original MFs.
- MFs area similarity (ρ): This metric measures the area similarity of the tuned MFs and the original ones.

Let us represent the definition parameters of the original and the tuned MF j as (a_j, b_j, c_j) and (a'_j, b'_j, c'_j) respectively, which can vary in their respective variation intervals $[I^l_{a_j}, I^r_{a_j}]$, $[I^l_{b_j}, I^r_{b_j}]$ and $[I^l_{c_j}, I^r_{c_j}]$, respectively. These intervals determine the maximum variation for each parameter and could be defined in a different way for different problems.

The δ metric can control the displacements in the central point of the MFs. It is based on computing the normalized distance between the central point of the tuned MF and the central point of the original MF, and it is calculated through obtaining the maximum displacement obtained in all the MFs. For each MF_j in the linguistic fuzzy partition, we define $\delta_j = |b_j - b'_j|/I$, where $I = (I^r_{b_j} - I^l_{b_j})/2$ represents the maximum variation for each central parameter. Thus δ^* is defined as $\delta^* = max_j\{\delta_j\}$ (the worst case). The δ^* metric takes values between 0 and 1, therefore values near to 1 show that the MFs present a great displacement. The following transformation is made so that this metric represents proximity (maximization):

$$\text{Maximize } \delta = 1 - \delta^* \ .$$

The γ metric is used to control the MF shapes. It is based on relating the left and right parts of the support of the original and the tuned MFs. Let us define $leftS_j = |a_j - b_j|$ as the amplitude of the left part of the original MF support and $rightS_j = |b_j - c_j|$ as the right part amplitude. Let us define $leftS'_j = |a'_j - b'_j|$ and $rightS'_j = |b'_j - c'_j|$ as the corresponding parts in the tuned MFs. γ_j is calculated using the following equation for each MF:

$$\gamma_j = \frac{min\{leftS_j/rightS_j\ ,\ leftS'_j/rightS'_j\}}{max\{leftS_j/rightS_j\ ,\ leftS'_j/rightS'_j\}} \ .$$

Values near to 1 mean that the left and right rate in the original MFs are highly maintained in the tuned MFs. Finally γ is calculated by obtaining the minimum value of γ_j (the worst case):

$$\text{Maximize } \gamma = min_j\{\gamma_j\} \ .$$

The ρ metric is used to control the area of the MF shapes. It is based on relating the areas of the original and the tuned MFs. Let us define A_j as the area of the triangle representing the original MF_j, and A'_j as the new area. ρ_j is calculated using the following equation for each MF:

$$\rho_j = \frac{min\{A_j, A'_j\}}{max\{A_j, A'_j\}} \ .$$

Values near to 1 mean that the original area and the tuned area of the MFs are more similar (less changes). The ρ metric is calculated by obtaining the minimum value of ρ_j (the worst case):

$$\text{Maximize } \rho = min_j\{\rho_j\} \ . \tag{1}$$

They propose a particular MOEA for regression problems with three objectives: minimization of the MSE, minimization of the number of rules (for which it is also included in the C_1 quadrant) and maximization of the semantic interpretability index. This MOEA is designed to generate a set of FRBSs with different trade-offs among accuracy, complexity and semantic interpretability, which allows the selection of the most appropriate solution from the final Pareto front depending on the expert preferences. The authors include examples that show how the change in the MFs has been almost imperceptible but involving improvements of 30% in accuracy.

7. Conclusions

In this contribution, we have presented a review of interpretability of fuzzy systems focused on the framework of linguistic FRBSs. A complete list of works on the use or the proposal of techniques or measures to take into account the interpretability of linguistic FRBSs, as a part of the problem of finding a good trade-off between interpretability and accuracy, have been studied in this paper. To this end, we have proposed a taxonomy with four quadrants (complexity or semantic interpretability at the level of RB or fuzzy partitions) as a way of organizing the different measures or constraints that we find in the literature to control interpretability in this area. We have analyzed the different measures proposed competing in the different quadrants. Since the interpretability of linguistic FRBSs is still an open problem, this will help researchers in this field to determine the most appropriate measure according to the part of the KB in which they want to maintain / improve the interpretability .

After studying the different works in the said topic, we can state that there is no single comprehensive measure to quantify the interpretability of linguistic models. In our opinion, to get a good global measure it would be necessary to consider appropriate measures from all the quadrants, in order to take into account together the different interpretability properties required for these kinds of systems. The different measures, from each quadrant, could be optimized as different objectives within a multi-objective framework. This would allow a search for a compromise among the different measures to take place taking accuracy into account. The main problem is that nowadays multi-objective optimization algorithms are not able to handle much more than three objectives adequately. Therefore, it is also necessary to find a way to combine them into a single index using weights or appropriate aggregation operators in order to give appropriate importance to one or another measure. One possibility is to aggregate complexity-based and sematic-based measures separately, giving way to two different indexes. This would allow the different trade-offs among accuracy, complexity and semantic interpretability to be found. In this sense, and taking into account the studied works in the different quadrants, we can make the following statements:

- In the quadrants C_1 and C_2, there are well-known and used measures to quantify complexity. These measures are widely accepted as the number of rules, number of conditions and number of features. Moreover, these measures are easy to use in practice since, for example, measuring the total number of conditions is a way to also take into account the remaining ones (such as fewer conditions in a model, smaller number of rules and/or smaller number of features).

- By contrast, there is no agreement in the choice of an appropriate measure in the C_3 and C_4 quadrants. Nevertheless, we can find two interesting ways of working for each quadrant respectively: On one hand, we consider that the use of the number of rules fired at the same time [21, 59] is a very promising measure or a way of working for C_3 if it is properly combined or adapted in order to also consider consistency of the rules. On the other hand, the use of a relative measure such as $GM3M$ [11] defined as a global semantic interpretability index to quantify interpretability with respect to the preferred reference fuzzy partitions that an expert could provide, seems a first solid attempt for measuring/comparing the models for C_4.

An interesting way to analyze how we could combine the measures of the different quadrants is to pay attention to what users and/or experts would consider interpretable, for instance, by using a web poll like in [8]. Alonso *et al* in [8] propose to use the knowledge extracted by means of a web poll of researchers familiarized and not familiarized with LFM. This poll is dedicated to determine how different people assess interpretability giving priority to different criteria.

Finally, we must point out that it is necessary to establish the measures for the different quadrants and, with respect to the aggregation of the different measures in a global index, the way to combine the measures by selecting the appropriate aggregation operators is not a trivial but an essential task.

References

[1] E. H. Mamdani, Application of fuzzy algorithms for control of simple dynamic plant, in: Proc. IEEE, Vol. 121, 1974, pp. 1585–1588.

[2] E. H. Mamdani, S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, International Journal of Man-Machine Studies 7 (1975) 1–13.

[3] T. Takagi, M. Sugeno, Fuzzy identification of systems and its application to modeling and control, IEEE Trans. Syst., Man, Cybern. 15 (1) (1985) 116–132.

[4] J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Accuracy improvements in linguistic fuzzy modeling, Vol. 129 of Studies in Fuzziness and Soft Computing, Springer, 2003.
[5] J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Interpretability issues in fuzzy modeling, Vol. 128 of Studies in Fuzziness and Soft Computing, Springer, 2003.
[6] S. M. Zhou, J. Q. Gan, Low-level interpretability and high-level interpretability: A unified view of data-driven interpretable fuzzy system modelling, Fuzzy Sets and Systems 159 (23) (2008) 3091–3131.
[7] C. Mencar, A. Fanelli, Interpretability constraints for fuzzy information granulation, Information Sciences 178 (24) (2008) 4585–4618.
[8] J. M. Alonso, L. Magdalena, G. G. Rodríguez, Looking for a good fuzzy system interpretability index: An experimental approach, International Journal of Approximate Reasoning 51 (2009) 115–134.
[9] R. Alcalá, M. J. Gacto, F. Herrera, J. Alcalá-Fdez, A multi-objective genetic algorithm for tuning and rule selection to obtain accurate and compact linguistic fuzzy rule-based systems, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 15 (5) (2007) 539–557.
[10] J. M. Alonso, L. Magdalena, S. Guillaume, Hilk: A new methodology for designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism, International Journal of Intelligent Systems 23 (7) (2008) 761–794.
[11] M. Gacto, R. Alcalá, F. Herrera, Integration of an index to preserve the semantic interpretability in the multi-objective evolutionary rule selection and tuning of linguistic fuzzy systems, IEEE Transactions on Fuzzy Systems 18 (3) (2010) 515–531.
[12] M. J. Gacto, R. Alcalá, F. Herrera, Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems, Soft Computing 13 (5) (2009) 419–436.
[13] 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 89 (2) (1997) 135–150.
[14] H. Ishibuchi, T. Nakashima, T. Murata, Three-objective genetics-based machine learning for linguistic rule extraction, Information Sciences 136 (2001) 109–133.
[15] H. Ishibuchi, Y. Nojima, Analysis of interpretability-accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning, International Journal of Approximate Reasoning 44 (1) (2007) 4–31.
[16] H. Ishibuchi, K. Nozaki, N. Yamamoto, H. Tanaka, Selecting fuzzy if-then rules for classification problems using genetic algorithms, IEEE Trans. Fuzzy Syst. 3 (3) (1995) 260–270.
[17] H. Ishibuchi, T. Yamamoto, Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining, Fuzzy Sets and Systems 141 (1) (2004) 59–88.
[18] P. Pulkkinen, J. Hytönen, H. Koivisto, Developing a bioaerosol detector using hybrid genetic fuzzy systems, Engineering Applications of Artificial Intelligence 21 (8) (2008) 1330–1346.
[19] U. Bodenhofer, P. Bauer, A formal model of interpretability of linguistic variables, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Interpretability issues in fuzzy modeling, Springer-Verlag, 2003, pp. 524–545.
[20] A. Botta, B. Lazzerini, F. Marcelloni, D. C. Stefanescu, Context adaptation of fuzzy systems through a multi-objective evolutionary approach based on a novel interpretability index, Soft Computing 13 (5) (2009) 437–449.
[21] F. Cheong, R. Lai, Constraining the optimization of a fuzzy logic controller using an enhanced genetic algorithm, IEEE Trans. Syst., Man, Cybern. - Part B: Cybernetics 30 (1) (2000) 31–46.
[22] J. Espinosa, J. Vandewalle, Constructing fuzzy models with linguistic integrity from numerical data-AFRELI algorithm, IEEE Trans. Fuzzy Syst. 8 (5) (2000) 591–600.
[23] P. Fazendeiro, J. V. de Oliveira, W. Pedrycz, A multiobjective design of a patient and anaesthetist-friendly neuromuscular blockade controller, IEEE Trans. on Biomedical Engineering 54 (9) (2007) 1667–1678.
[24] Y. Jin, W. von Seelen, B. Sendhoff, On generating fc^3 fuzzy rule systems from data using evolution strategies, IEEE Transactions on Systems, Man, and Cybernetics 29 (6) (1999) 829–845.
[25] C. Mencar, G. Castellano, A. M. Fanelli, Distinguishability quantification of fuzzy sets, Information Sciences 177 (2007) 130–149.
[26] D. Nauck, Measuring interpretability in rule-based classification systems, in: Proceedings of the 12th IEEE International Conference on Fuzzy Systems, Vol. 1, 2003, pp. 196–201.
[27] J. V. de Oliveira, Semantic constraints for membership function optimization, IEEE Trans. Syst., Man, Cybern. - Part A: Systems and Humans 29 (1) (1999) 128–138.
[28] J. V. de Oliveira, Towards neuro-linguistic modeling: constraints for optimization of membership functions, Fuzzy Sets and Systems 106 (3) (1999) 357–380.
[29] S. Guillaume, B. Charnomordic, A new method for inducing a set of interpretable fuzzy partitions and fuzzy inference systems from data, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Interpretability issues in fuzzy modeling, Springer-Verlag, 2003, pp. 148–175.
[30] S. Guillaume, B. Charnomordic, Generating an interpretable family of fuzzy partitions from data, IEEE Transactions on Fuzzy Systems 12 (3) (2004) 324–335.
[31] F. Liu, C. Quek, G. S. Ng, A novel generic hebbian ordering-based fuzzy rule base reduction approach to mamdani neuro-fuzzy system, Neural Computation 19 (2007) 1656–1680.
[32] H. Ishibuchi, T. Murata, Multi-objective genetic local search algorithm, in: Proc. third IEEE International Conference on Evolutionary Computation, Japan, 1996, pp. 119–124.
[33] C. A. Pena-Reyes, M. Sipper, Fuzzy CoCo: Balancing accuracy and interpretability of fuzzy models by means of coevolution, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Accuracy improvements in linguistic fuzzy modeling, Springer-Verlag, 2003, pp. 119–146.
[34] H. Ishibuchi, T. Murata, I. B. Turksen, Selecting linguistic classification rules by two-objective genetic algorithms, in: Proc. of the IEEE International Conference on Syst., Man, Cybern., Vancouver, Canada, 1995, pp. 1410–1415.
[35] Y. Jin, Fuzzy modeling of high-dimensional systems: Complexity reduction and interpretability improvement, IEEE Transactions on Fuzzy Systems 8 (2) (2000) 212–221.
[36] S. Guillaume, Designing fuzzy inference systems from data: An interpretability-oriented review, IEEE Transactions on Fuzzy Systems 9 (3) (2001) 426–443.

[37] H. Ishibuchi, T. Yamamoto, Trade-off between the number of fuzzy rules and their classification performance, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Accuracy improvements in linguistic fuzzy modeling, Springer-Verlag, 2003, pp. 72–99.
[38] H. Ishibuchi, T. Yamamoto, Interpretability issues in fuzzy genetics-based machine learning for linguistic modelling, in: J. Lawry, J. G. Shanahan, A. L. Ralescu (Eds.), Modelling with words: Learning, fusion, and reasoning within a formal liguistic representation framework, Vol. 2873 of Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, 2003, pp. 209–228.
[39] O. Cordón, F. Herrera, L. Magdalena, P. Villar, A genetic learning process for the scaling factors, granularity and contexts of the fuzzy rule-based system data base, Information Science 136 (2001) 85–107.
[40] O. Cordón, F. Herrera, P. Villar, Generating the knowledge base of a fuzzy rule-based system by the genetic learning of data base, IEEE Transactions on Fuzzy Systems 9 (4) (2001) 667–674.
[41] R. Alcalá, J. Alcalá-Fdez, F. Herrera, J. Otero, Genetic learning of accurate and compact fuzzy rule based systems based on the 2-tuples linguistic representation, International Journal of Approximate Reasoning 44 (1) (2007) 45–64.
[42] F. Herrera, L. Martínez, A 2-tuple fuzzy linguistic representation model for computing with words, IEEE Trans. Fuzzy Syst. 8 (6) (2000) 746–752.
[43] O. Cordón, M. del Jesus, F. Herrera, L. Magdalena, P. Villar, A multiobjective genetic learning process for joint feature selection and granularity and context learning in fuzzy rule-based classification systems, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Interpretability issues in fuzzy modeling, Springer-Verlag, 2003, pp. 79–99.
[44] J. Casillas, O. Cordón, F. Herrera, L. Magdalena, Interpretability improvements to find the balance interpretability-accuracy in fuzzy modeling: an overview, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Interpretability issues in fuzzy modeling, Springer-Verlag, 2003, pp. 3–22.
[45] R. Mikut, J. Jakel, L. Grall, Interpretability issues in data-based learning of fuzzy systems, Fuzzy Sets and Systems 150 (2005) 179–197.
[46] J. Casillas, O. Cordón, M. J. del Jesus, F. Herrera, Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction, IEEE Trans. Fuzzy Syst. 13 (1) (2005) 13–29.
[47] K. Narukawa, Y. Nojima, H. Ishibuchi, Modification of evolutionary multiobjective optimization algorithms for multiobjective design of fuzzy rule-based classification systems, in: Proc. of the 2005 IEEE International Conference on Fuzzy Systems, Reno, USA, 2005, pp. 809–814.
[48] K. Deb, A. Pratab, S. Agrawal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evolutionary Computation 6 (2) (2002) 182–197.
[49] R. Alcalá, J. Alcalá-Fdez, F. Herrera, A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection, IEEE Trans. Fuzzy Syst. 15 (4) (2007) 616–635.
[50] P. Pulkkinen, H. Koivisto, Fuzzy classifier identification using decision tree and multiobjective evolutionary algorithms, International Journal of Approximate Reasoning 48 (2) (2008) 526–543.
[51] F. Herrera, Genetic fuzzy systems: taxonomy, current research trends and prospects, Evolutionary Intelligence 1 (2008) 27–46.
[52] J. Quinlan (Ed.), C4.5: Programs for Machine Learning, Morgan Kaufman Publishers, 1993.
[53] R. Alcalá, P. Ducange, F. Herrera, B. Lazzerini, F. Marcelloni, A multi-objective evolutionary approach to concurrently learn rule and data bases of linguistic fuzzy rule-based systems, IEEE Transactions on Fuzzy Systems 17 (5) (2009) 1106–1122.
[54] P. Pulkkinen, H. Koivisto, A dynamically constrained multiobjective genetic fuzzy system for regression problems, IEEE Transactions on Fuzzy Systems 18 (1) (2010) 161–177.
[55] M. Setnes, R. Babuška, U. Kaymak, H. R. van Nauta Lemke, Similarity measures in fuzzy rule base simplification, IEEE Trans. Syst., Man, Cybern. - Part B: Cybernetics 28 (3) (1998) 376–386.
[56] D. Tikk, T. Gedeon, K. Wong, A feature ranking algorithm for fuzzy modelling problems, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Interpretability issues in fuzzy modeling, Springer-Verlag, 2003, pp. 176–192.
[57] P. A. Devijver, J. Kittler, Pattern Recognition: A Statistical Approach, Prentice Hall, London, 1982.
[58] V. Vanhoucke, R. Silipo, Interpretability in multidimensional classification, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Interpretability issues in fuzzy modeling, Springer-Verlag, 2003, pp. 193–217.
[59] F. Cheong, R. Lai, Constrained optimization of genetic fuzzy systems, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Accuracy Improvements in Linguistic Fuzzy Modeling, Springer-Verlag, 2003, pp. 46–71.
[60] Y. Jin, Generating distinguishable, complete, consistent and compact fuzzy systems using evolutionary algorithms, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Accuracy improvements in linguistic fuzzy modeling, Springer-Verlag, 2003, pp. 100–118.
[61] W. Pedrycz, Expressing relevance interpretability and accuracy of rule based systems, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Interpretability issues in fuzzy modeling, Springer-Verlag, 2003, pp. 546–567.
[62] W. Pedrycz, J. de Oliveira, Optimization of fuzzy models, IEEE Trans. Systems Man Cybernet. Part B 26 (4) (1996) 627–636.
[63] T. Furuhashi, T. Suzuki, On interpretability of fuzzy models based on conciseness measure, in: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'01), 2001, pp. 284–287.
[64] T. Suzuki, T. Kodama, T. Furuhashi, H. Tsutsui, Fuzzy modeling using genetic algorithms with fuzzy entropy as conciseness measure, Information Sciences 136 (1-4) (2001) 53–67.
[65] T. Suzuki, T. Furuhashi, Conciseness of fuzzy models, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Interpretability issues in fuzzy modeling, Springer-Verlag, 2003, pp. 568–586.
[66] S. T. A. De Luca, A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory, Information and Control 20 (1972) 301–312.
[67] P. Fazendeiro, J. V. de Oliveira, A working hypothesis on the semantics/accuracy synergy, in: EUSFLAT-LFA 2005, 2005, pp. 266–271.
[68] J. V. de Oliveira, P. Fazendeiro, On the achievement of both accurate and interpretable fuzzy systems using data-driven design processes, in: J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Accuracy improvements in linguistic fuzzy modeling, Springer-Verlag, 2003, pp. 147–162.

3. Algoritmos Evolutivos Multi-Objetivo que Combinan las Técnicas de Ajuste y de Selección de Reglas para Obtener Sistemas Basados en Reglas Difusas Lingüísticos Precisos y Compactos - *Multi-objective Genetic Algorithms for Tuning and Rule Selection to Obtain Accurate and Compact Linguistic Fuzzy Rule-Based Systems*

Las publicaciones en revista asociadas a esta parte son:

3.1. *Propuesta Inicial de un Algoritmo Evolutivo Multi-Objetivo para Problemas de Regresión*

- R. Alcalá, M.J. Gacto, F. Herrera, J. Alcalá-Fdez, A Multi-objective Genetic Algorithm for Tuning and Rule Selection to Obtain Accurate and Compact Linguistic Fuzzy Rule-Based Systems. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 15:5 (2007) 539-557, doi:10.1142/S0218488507004868.
    - Estado: Publicado.
    - Índice de Impacto (JCR 2007): 0,376.
    - Área de Conocimiento: Computer Science, Artificial Intelligence. Ranking 81 / 93.
    - Citas: 17.

International Journal of Uncertainty,
Fuzziness and Knowledge-Based Systems
Vol. 15, No. 5 (2007) 539–557

A MULTI-OBJECTIVE GENETIC ALGORITHM FOR TUNING AND RULE SELECTION TO OBTAIN ACCURATE AND COMPACT LINGUISTIC FUZZY RULE-BASED SYSTEMS*

R. ALCALÁ*, M. J. GACTO† and F. HERRERA‡
Dept. Computer Science and A.I., University of Granada,
Granada, E-18071, Spain,
* *alcala@decsai.ugr.es*
† *herrera@decsai.ugr.es*
‡ *mjgacto@ugr.es*

J. ALCALÁ-FDEZ
Dept. Computer Science, University of Jaen, Jaen, E-23071, Spain
jalcala@ujaen.es

Received 2 January 2007
Revised 18 August 2007

This work proposes the application of Multi-Objective Genetic Algorithms to obtain Fuzzy Rule-Based Systems with a better trade-off between interpretability and accuracy in linguistic fuzzy modelling problems. To do that, we present a new post-processing method that by considering selection of rules together with tuning of membership functions gets solutions only in the Pareto zone with the highest accuracy, i.e., containing solutions with the least number of possible rules but still presenting high accuracy. This method is based on the well-known SPEA2 algorithm, applying appropriate genetic operators and including some modifications to concentrate the search in the desired Pareto zone.

Keywords: Multi-Objective Genetic Algorithms; Linguistic Modelling; Interpretability-Accuracy Trade-Off; Rule Selection; Tuning of Membership Functions.

1. Introduction

One of the aims in focusing the research in the Linguistic Fuzzy Modelling area in recent years is the trade-off between interpretability and accuracy.[1] Of course, the ideal thing would be to satisfy both criteria to a high degree, but since they are contradictory issues generally it is not possible.

*Supported by the Spanish Project TIN-2005-08386-C05-01.

A widely-used approach to improve the accuracy of linguistic Fuzzy Rule-Based Systems (FRBSs) is the *tuning* of Membership Functions (MFs),[1–10] which refines a previous definition of the Data Base (DB) once the rule base has been obtained. Although tuning usually improves the system performance, sometimes a large number of rules is used to reach an acceptable degree of accuracy. In this case, some works[1,5] consider the selection of rules together with the tuning of MFs but only considering performance criteria.

In this contribution, we focus on this problem by using Genetic Algorithms as a tool for evolving the MFs parameters and rule base size and by coding all of them (rules and parameters) in the same chromosome. Since the problem presents multi-objective nature we could consider the use of Multi-Objective Genetic Algorithms (MOGAs)[11–15] to obtain a set of solutions with different degrees of accuracy and number of rules by using both characteristics as objectives.

Although there are some works in the literature using MOGAs to improve the difficult trade-off between interpretability and accuracy of FRBSs,[16–25] practically all these works were applied to classification problems trying to obtain the complete Pareto (set of non-dominated solutions with different trade-off) by selecting or learning the set of rules better representing the example data, i.e., improving the system classification ability and decreasing the system complexity but not considering learning or tuning of the fuzzy system parameters (which involves another type of Pareto front, a more complicated search space and therefore needs different considerations respect to the works in the existing literature).

In this way, our main interest is to design an appropriate MOGA for this type of problem due to the fact that standard MOGAs can present some problems. As said, MOGAs are generally based on obtaining a set of non-dominated solutions. However, in this case, there are solutions that are not interesting although they are in the Pareto frontier. For example, non-dominated solutions with a small number of rules and high error are not interesting since they have not the desired trade-off between accuracy and interpretability. Furthermore, the existence of these kinds of solutions favours the selection of solutions with very different number of rules and accuracy to apply the crossover operator, which gives results with poor accuracy (the tuning parameters would be very different and the crossover would not have any sense except for exploring new combinations of rules).

In our proposal, we concentrate the search in the Pareto zone with still accurate solutions trying to obtain the least number of possible rules. To do that, we propose a modification of the well-known SPEA2[26] (Strength Pareto Evolutionary Algorithm 2) that considering the rule selection together with the tuning of MFs concentrates the search in the Pareto zone having accurate solutions with the least number of possible rules, the Accuracy-Oriented SPEA2 (SPEA2$_{ACC}$). Besides, we have performed the same modification and experiments with NSGA-II[27] (Nondominated Sorting Genetic Algorithm II), showing that this approach is not the most adequate for this problem.

This paper is arranged as follows. First, a brief summary of different proposals to improve the balance between interpretability and accuracy is presented, specially taking into account those considering MOGAs for this purpose. In section 3, we present a study of the estimated Pareto frontier for this problem (tuning and rule selection). SPEA2$_{ACC}$ algorithm is introduced in Section 4 together with the modifications proposed on SPEA2 and the genetic operators considered. Section 5 shows an experimental study of the proposed methods in a real-world problem. Finally, Section 6 gives some conclusions.

2. Interpretability-Accuracy Trade-off of FRBSs

Fuzzy Modelling (FM) usually comes with two contradictory requirements to the obtained model: the *interpretability*, capability to express the behaviour of the real system in an understandable way, and the *accuracy*, capability to faithfully represent the real system. Since they are contradictory issues, more priority has generally been given to one of them (defined by the problem nature), leaving the other one in the background. Two FM approaches arise depending on the main objective to be considered:

- *Linguistic FM*, mainly developed by means of linguistic (or Mamdani) FRBSs,[28,29] which is focused on the interpretability.
- *Precise FM*, mainly developed by means of Takagi-Sugeno FRBSs,[30] which is focused on the accuracy.

Regardless of the approach, a common scheme has been considered to attain the desired balance between interpretability and accuracy (Figure 1 graphically shows this operation mode):

(1) Firstly, the main objective (interpretability or accuracy) is tackled defining a specific model structure to be used, thus setting the FM approach.

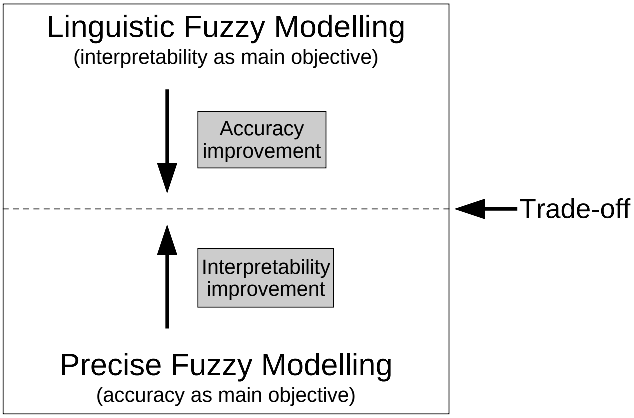


Fig. 1. Improvements of interpretability and accuracy in fuzzy modelling.

(2) Then, the modelling components (the model structure and/or the modelling process) are improved by means of different mechanisms to compensate for the initial difference between both requirements. Thus, accuracy improvements are proposed in linguistic FM at the cost of part of the interpretability whilst interpretability improvements are proposed in precise FM at the cost of part of the accuracy.

Actually, the interpretability-accuracy trade-off is a very important branch of research nowadays.[1,31] Focusing on *Linguistic FM with improved accuracy*[1] (still nearer of the interpretability) we can find many examples in the existing literature. This approach has been performed by learning/tuning the MFs by defining their parameters or shapes,[2–10] their types (triangular, trapezoidal, etc.),[32] or their context (defining the whole semantics),[5,33,34] learning the granularity (number of linguistic terms) of the fuzzy partitions,[33,35] or extending the model structure by using linguistic modifiers,[5,36,37] weights (importance factors for each rule),[2,38–40] or hierarchical architectures (mixing rules with different granularities),[38,42] among others. The main problem of these approaches is that although the system accuracy can be greatly improved (e.g., with a simple tuning of MFs), the original interpretability of the linguistic models is lost to some degree giving way to more complex systems or less interpretable rule structures.

Additionally, although rule base reduction[5,41,42] and input variable selection[43,44] processes improve the interpretability, they can also help to improve the accuracy when redundancy and inconsistency criteria are considered (but usually these improvements are not very significant).

Within the framework of linguistic FM (without improved accuracy), a new and most recent possibility is the use of Multi-Objective Evolutionary Algorithms (MOEAs)[11–15] to improve the difficult trade-off between interpretability and accuracy of FRBSs, considering different performance and interpretability measures as objectives.[16–25] Since this problem presents a multi-objective nature the use of these kinds of algorithms to obtain a set of solutions with different degrees of accuracy and interpretability is an interesting way to work. All of the works in this recent topic only consider quantitative measures of the system complexity (number of rules, number of characteristics in the antecedents, etc.) in order to improve the interpretability of such systems, rarely considering qualitative measures. Furthermore, we can point out that practically all these methods were applied to classification problems for rule selection or rule learning, without considering learning or tuning the MFs or more flexible rule representations, i.e., performing Linguistic FM with improved interpretability to obtain a set of solutions with different trade-offs but nearer the interpretability than the accuracy.

In this way, our main aim in this contribution will be to attain the desired balance by maintaining the improved accuracy that a tuning of MFs could give but trying to obtain more compact models by using MOGAs if it is possible, i.e., to obtain simpler and still accurate linguistic fuzzy models by also considering a tuning of the system parameters. This way to work represents a more complex search space and therefore needs a deeper analysis of the Pareto frontier and different considerations respect to the MOGAs in the existing literature.

3. Interpretability-Accuracy Pareto Frontier by Selecting Rules and Tuning Membership Functions

In this section, we present a study of the kinds of solutions we could find in the optimal Pareto frontier when the system error and the number of rules (both considered as objectives) are optimized by tuning the MFs and selecting the most promising rules. In this way, we can obtain an approximation of the optimal Pareto that can help to determine the desired Pareto zone.

Tuning of MFs usually needs an initial model with large number of rules to get an appropriate level of accuracy. Generally, to obtain a good number of initial rules, methods ensuring covering levels higher than needed are used. In this way, we could obtain rules that being needed at first could be unnecessary once the tuning is applied or rules that could impede the tuning of the remaining ones in order to obtain the global optimum in terms of the accuracy (better configuration of rules to get the minimum error after tuning of the parameters). Thus, we can find the following types of rules respect to this global optimum in the complete set of rules: *Bad Rules* (erroneous or conflicting rules) that degrade the system performance (rules that are not included in the most accurate final solution); *Redundant or Irrelevant Rules* that do not significantly improve the system performance; *Complementary Rules* that complement some others slightly improving the system performance; and *Important Rules* that should not be removed to obtain a reasonable system performance. Obviously, this is a simplification of the problem by only considering in principle the most accurate solution in order to have an idea of the shape of the optimum Pareto. On the other hand, to determine those types of rules in advance is impossible since it directly depends on each concrete configuration of rules and still more on the optimal configuration of the MF parameters for each rule configuration. Therefore, this is impossible to establish any criteria that could be used in the search process.

However, by taking into account the possible existence of these kinds of rules, different rule configurations and different tuning parameters, we can estimate the following zones in the space of the objectives:

- Zone with Bad Rules, which contains solutions with bad rules. In this zone, the Pareto front does not exist given that removing these kinds of rules would improve the accuracy and these solutions would be dominated by others.
- Zone with Redundant or Irrelevant Rules, which is comprised of solutions without bad rules but still maintaining redundant or irrelevant rules. By deleting these kinds of rules the accuracy would be practically the same.
- Zone with Complementary Rules, comprised of solutions without any bad or redundant rule. By removing these rules the accuracy would be slightly decreased.

- Zone with Important Rules, which contains solutions only comprised of essential rules. By removing these kinds of rules the accuracy is really affected.

In Figure 2, we can find an approximation of the optimal Pareto in the problem of tuning and rule selection with the double objective of simplicity and accuracy. This figure shows the different zones in the space of the objectives together with the desired Pareto zone to find solutions with good trade-off. This zone corresponds to the zone of complementary rules, i.e., we would like to delete all the possible rules but without seriously affecting the accuracy of the model finally obtained.

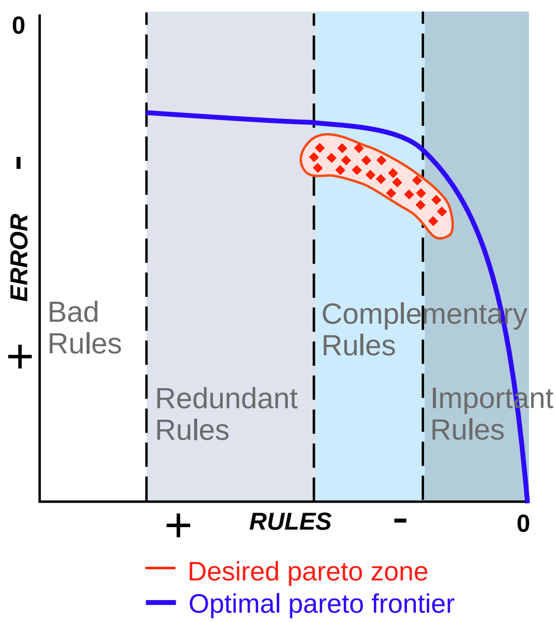


Fig. 2. Estimation of the pareto frontier considering rule selection and tuning of parameters.

Taking into account what we previously exposed, the MOGA should not obtain all the Pareto front since it is difficult to obtain accurate solutions by favouring the crossing of solutions with very different rule configurations (those in the Pareto), which try to obtain the optimum by learning very different parameters for the MFs. In the next section, we present a modification of SPEA2[26] with the main aim of guiding the search towards the desired zone.

4. A Proposal to Evolve Accuracy-Oriented Pareto Sets: the SPEA2$_{ACC}$ Algorithm

This section presents a new algorithm to get solutions with high accuracy and the least possible number of rules by performing rule selection together with a tuning of the MF parameters. In this way, since this algorithm is based on the well known SPEA2[26] we firstly introduce the basis of this algorithm. Then we describe the changes for guiding the search towards the desired Pareto zone and the main components needed to apply this algorithm to this specific problem: the coding scheme and the genetic operators.

4.1. ***SPEA2 Basis***

The SPEA2 algorithm[26] (*Strength Pareto Evolutionary Algorithm 2 for multi-objetive optimization*) is one of the most used techniques for solving problems with multi-objective nature. This algorithm was designed to overcome the problems of its predecessor, the SPEA algorithm.[15] In contrast with SPEA, SPEA2: (1) incorporates a fine-grained fitness assignment strategy which takes into account for each individual the number of individuals that it dominates and the number of individuals by which it is dominated; (2) uses the nearest neighbour density estimation technique which guides the search more efficiently; (3) has an enhanced archive truncation method which guarantees the preservation of boundary solutions. Next, we briefly describe the complete SPEA2 algorithm.

SPEA2 uses a fixed population and archive size. The population forms the current base of possible solutions, while the archive contains the current solutions. The archive is constructed and updated by copying all non-dominated individuals in both archive and population into a temporary archive. If the size of this temporary archive differs from the desired archive size, individuals are either removed or added as necessary. Individuals are added by selecting the best dominated individuals, while the removal process uses a heuristic clustering routine in the objective space. The motivation for this is that one would like to try to ensure that the archive contents represent distinct parts of the objective space.

The fitness assignment strategy takes into account both dominating and dominated solutions for each individual. Let P_t and $\overline{P}_t$ denote the population and the archive respectively, each individual i in $P_t + \overline{P}_t$ is assigned a strength value $S(i)$, the number of solutions it dominates,

$$S(i) = \| \{ j \mid j \in P_t + \overline{P}_t \ \wedge \ i \succ j \} \| \tag{1}$$

where $\| \cdot \|$ represents the cardinality of a set, $+$ stands for multiset union and the symbol $\succ$ corresponds to the Pareto dominance relation. Based on the value of $S(i)$, a raw fitness value, $R(i)$, is given to the individual i,

$$R(i) = \sum_{j \in P_t + \overline{P}_t,\ j \succ i} S(j) \tag{2}$$

It is important to notice that fitness is to be minimized here, i.e., $R(i) = 0$ corresponds to a nondominated individual, while a high $R(i)$ value means that i is dominated by many individuals (which in turn dominate many other individuals). This scheme is illustrated in Figure 3. The final fitness value is assigned by adding a

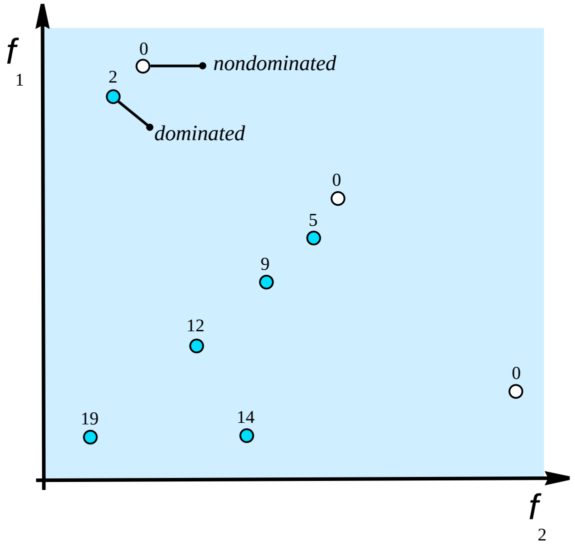


Fig. 3. The raw SPEA2 fitness values for a maximization problem with two objectives f_1 and f_2.

density value. The density function value, $D(i)$, is estimated in the objective space,

$$D(i) = \frac{1}{\delta_i^k + 2} \tag{3}$$

where δ_i^k denotes the k-th nearest distance for the ith individual among P_t and $\overline{P}_t$ in objective space. k is usually set as $\sqrt{N + \overline{N}}$ truncated to an integer, where N is the population size and $\overline{N}$ the archive size. Finally, the fitness value for the i-th individual is calculated as,

$$F(i) = R(i) + D(i) \tag{4}$$

From the definition above, a better solution will be assigned a smaller fitness value. Finally, when selecting individuals for participating in the next generation all candidates are selected from the archive using a binary tournament selection scheme.

According to the descriptions of the authors in,[26] the outline of the SPEA2 algorithm is:

Input: N (population size),
$\overline{N}$ (external population size),
T (maximum number of generations).
Output: A (non-dominated set).

(1) Generate an initial population P_0 and create the empty external population $\overline{P}_0 = \emptyset$.
(2) Calculate fitness values of individuals in P_t and $\overline{P}_t$.

(3) Copy all non-dominated individuals in $P_t \cup \overline{P}_t$ to $\overline{P}_{t+1}$. If $|\overline{P}_{t+1}| > \overline{N}$ apply truncation operator. If $|\overline{P}_{t+1}| < \overline{N}$ fill with dominated in $P_t \cup \overline{P}_t$.
(4) If $t \geq T$, return A and stop.
(5) Perform binary tournament selection with replacement on $\overline{P}_{t+1}$ in order to fill the mating pool.
(6) Apply recombination and mutation operators to the mating pool and set P_{t+1} to the resulting population. Go to step 2 with $t = t + 1$.

4.2. *The SPEA2$_{ACC}$ algorithm*

In the following, the main aspects and components needed to design the proposed algorithm are explained. They are:

- Modifications Applied on SPEA2 to guide the search.
- Coding scheme and initial gene pool.
- Objectives considered for chromosome evaluation.
- Crossover and mutation operators.

4.2.1. *Modifications applied on SPEA2*

In order to focus the search on the desired Pareto zone, high accuracy with least possible number of rules, we propose two main changes on the SPEA2 algorithm with the aim of giving more selective pressure to those solutions that have a high accuracy. The proposed changes are described next:

- A restarting operator is applied exactly at the mid of the algorithm, by maintaining the most accurate individual as the sole individual in the external population ($\overline{P}_{t+1}$ with size 1) and obtaining the remaining individuals in the population (P_{t+1}) with the same rule configuration of the best individual and tuning parameters generated at random within the corresponding variation intervals. This operation is performed in step 4 then returning to step 2 with $t = t+1$. In this way, we concentrate the search only in the desired pareto zone (similar solutions in a zone with high accuracy).
- In each stage of the algorithm (before and after restarting), the number of solutions in the external population ($\overline{P}_{t+1}$) considered to form the mating pool is progressively reduced, by focusing only on those with the best accuracy. To do that, the solutions are sorted from the best to the worst (considering accuracy as sorting criterion) and the number of solutions considered for selection is reduced progressively from 100% at the beginning to 50% at the end of each stage.

Besides, we have to highlight that the way to create the solutions of the initial population for the part of rule selection is a very important factor. Usually, a Genetic Algorithm generates the initial population totally at random (random selection of the initial rules). However, in this case, to get solutions with a high accuracy we should not lose rules that could present a positive cooperation once their MF parameters have been evolved. The best way to do this is to start with solutions selecting all the possible rules which favours a progressive extraction of bad rules (those that do not improve with the tuning of parameters), only by means of the mutation at the beginning and then by means of the crossover.

4.2.2. *Coding scheme and initial gene pool*

A double coding scheme for both *rule selection* (C_S) and *tuning* (C_T) is used:

$$C^p = C_S^p C_T^p$$

- For the C_S part, the coding scheme consists of binary-coded strings with size m (with m being the number of initial rules). Depending on whether a rule is selected or not, values '1' or '0' are respectively assigned to the corresponding gene.

$$C_S^p = (c_{S1}, \ldots, c_{Sm}) \mid c_{Si} \in \{0, 1\} \ .$$

- For the C_T part, a real coding is considered, being m^i the number of labels of each of the n variables comprising the DB.

$$C_i = (a_1^i, b_1^i, c_1^i, \ldots, a_{m^i}^i, b_{m^i}^i, c_{m^i}^i), \ \ i = 1, \ldots, n \ ,$$
$$C_T^p = C_1 C_2 \ldots C_n \ .$$

The initial population is obtained in the following way:

(1) For the C_T part the initial DB is included as first individual. The remaining individuals are generated at random within the corresponding variation intervals. Such intervals are calculated from the initial DB. For each MF $C_i^j = (a^j, b^j, c^j)$, the variation intervals are calculated in the following way:

$$[I_{a^j}^l, I_{a^j}^r] = [a^j - (b^j - a^j)/2, a^j + (b^j - a^j)/2] \tag{5}$$

$$[I_{b^j}^l, I_{b^j}^r] = [b^j - (b^j - a^j)/2, b^j + (c^j - b^j)/2] \tag{6}$$

$$[I_{c^j}^l, I_{c^j}^r] = [c^j - (c^j - b^j)/2, c^j + (c^j - b^j)/2] \tag{7}$$

(2) For the C_S part all genes take value '1' in all the individuals of the initial population in order to favour a progressive extraction of the worst rules.

4.2.3. *Objectives*

Two objectives are minimized to get the desired trade-off: the number of rules (interpretability) and the Mean Squared Error (accuracy),

$$\text{MSE} = \frac{1}{2 \cdot |E|} \sum_{l=1}^{|E|} (F(x^l) - y^l)^2,$$

with $|E|$ being the size of a data set E, $F(x^l)$ being the output obtained from the FRBS decoded from the mentioned chromosome when the l-th example is considered and y^l being the known desired output. The fuzzy inference system considered to obtain $F(x^l)$ is the *centre of gravity weighted by the matching* strategy as defuzzification operator and the *minimum t-norm* as implication and conjunctive operators.

4.2.4. *Crossover and mutation operators*

The crossover operator depends on the chromosome part where it is applied:

- In the C_T part, the BLX-0.5[45] crossover is used. This operator is based on the the concept of environments (the offspring are generated around one parent). These kinds of operators present a good cooperation when they are introduced within evolutionary models forcing the convergence by pressure on the offspring. Figure 4 depicts the behaviour of this operator, which allow the offspring genes to be around a wide zone determined by both parent genes.

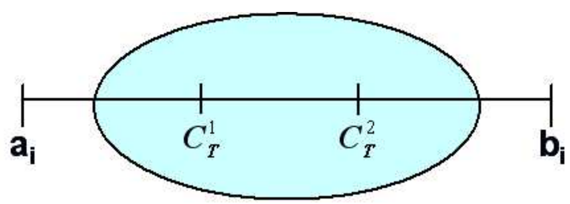


Fig. 4. Scheme of the behaviour of the BLX-α operator.

  The BLX is described as follows. Let us assume that $X = (x_1 \cdots x_n)$ and $Y = (y_1 \cdots y_n)$, $(x_i, y_i \in [a_i, b_i] \subset \Re, i = 1 \cdots n)$, are two real-coded chromosomes that are going to be crossed. The BLX operator (with $\alpha = 0.5$) generates one descendent $Z = (z_1, \cdots, z_g)$ where z_i is a randomly (uniformly) chosen number from the interval $[l_i, u_i]$, with $l_i = max\{a_i, c_{min} - I\}$, $u_i = min\{b_i, c_{max} + I\}$, $c_{min} = min\{x_i, y_i\}$, $c_{max} = max\{x_i, y_i\}$ and $I = (c_{max} - c_{min}) \cdot \alpha$.
- In the C_S part, the HUX[46] crossover is used. The HUX crossover exactly interchanges the mid of the alleles that are different in the parents (the genes to be crossed are randomly selected among those that are different in the parents). This operator ensures the maximum distance of the offspring to their parents (exploration). Figure 5 depicts the behaviour of this operator.

Finally, four offspring are generated by combining the two from the C_S part with the two from the C_T part (the two with the best accuracy are considered to be included in the population). The mutation operator changes a gene value at random in the C_S and C_T parts (one in each part) with probability P_m.

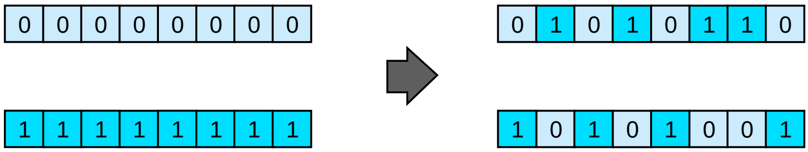


Fig. 5. Scheme of the behaviour of the HUX operator.

Table 1. Methods considered for comparison.

Ref.	Method	Description
47	**WM**	Wang & Mendel algorithm
5	**WM+T**	Tuning of Parameters
5	**WM+S**	Rule Selection
5	**WM+TS**	Tuning and Rule Selection
26	**SPEA2**	SPEA2 Algorithm
—	**SPEA2**$_{ACC}$	Accuracy-Oriented SPEA2
27	**NSGAII**	NSGA-II algorithm
—	**NSGAII**$_{ACC}$	Accuracy-Oriented NSGA-II

5. Experiments

To evaluate the usefulness of the method proposed, SPEA2$_{ACC}$, we have considered a real-world problem[49] with 4 input variables that consists of estimating the maintenance costs of medium voltage lines in a town. The methods considered for the experiments are briefly described in Table 1. WM[47] method is considered to obtain the initial rule base to be tuned. T and S methods perform the tuning of parameters and rule selection respectively. TS indicates tuning together with rule selection in the same algorithm. All of them consider the accuracy of the model as the sole objective. The remaining are MOGAs with and without the proposed modifications (all of them perform rule selection with tuning of parameters considering two objectives, accuracy and number of rules). However, we have to highlight that all of them consider the same population initialization, i.e., they start considering all the candidate rules for the initial individuals in order to see better the influence of the changes applied on the original SPEA2.

The linguistic partitions are comprised by *five linguistic terms* with triangular shape. The values of the input parameters considered by all the MOGAs studied are presented as follows: population size of 200, external population size of 61 (in the case of SPEA2 and SPEA2$_{ACC}$), 50,000 evaluations and 0.2 as mutation probability per chromosome.

5.1. *Problem description*

In Spain, electrical industries do not charge the energy bill directly to the final user, but they share the ownership of an enterprise (called R.E.E., Red Eléctrica

Table 2. Electrical problem characteristics.

Input variable X_1:	Sum of the lengths of all streets in the town
Input variable X_2:	Total area of the town
Input variable X_3:	Area that is occupied by buildings
Input variable X_4:	Energy supply to the town
Output variable Y:	Maintenance costs of the medium voltage lines
Number of examples:	1,059
Domain of X_1:	[0, 11]
Domain of X_2:	[0.15, 8.55]
Domain of X_3:	[1.64, 142.5]
Domain of X_4:	[1, 165]
Range of Y:	[0, 8546.03]

Española) which gets all payments and then distributes them according to some complex criteria (amount of power generation of every company, number of customers, etc.).

In the last years, some of these companies have asked for the rules to be revised. One of the proposed modifications involved a redistribution of the maintenance costs of the network. To compute the maintenance costs of town medium voltage lines, there is a need to know which would be the total line length if the installation made had been the optimal one. Clearly, it is impossible to obtain this value by directly measuring it, since the medium voltage lines existing in a town have been installed incrementally, according to its own electrical needs in each moment.

For this reason, the consideration of models becomes useful to compute the maintenance costs of the medium voltage electrical network in a town.[48,49] These estimations allow electrical companies to justify their expenses. Moreover, the model must be able to explain how a specific value is computed to a certain town. Our objective will be to relate the *maintenance costs of the medium voltage lines* with the following four variables: *sum of the lengths of all streets in the town*, *total area of the town*, *area that is occupied by buildings*, and *energy supply to the town*. We will deal with estimations of minimum maintenance costs based on a model of the optimal electrical network for a town in a sample of 1,059 towns. Table 2 presents a summary of the main characteristics of the problem.

To develop the different experiments, we consider a *5-folder cross-validation model*, i.e., 5 random partitions of data each with 20% (4 of them with 211 examples and one of them with 212 examples)[a], and the combination of 4 of them (80%) as training and the remaining one as test. For each one of the 5 data partitions, the tuning methods have been run 6 times, showing for each problem the average results of a total of 30 runs. In the case of methods with multi-objective approach, the averaged values are calculated considering the most accurate solution from each Pareto obtained. In this way, the multi-objective algorithms can be compared

[a]These data sets are available at: *http://decsai.ugr.es/∼casillas/fmlib/*.

Table 3. Results obtained by the studied methods.

Method	**#R**	**MSE**$_{tra}$	σ_{tra}	**t-test**	**MSE**$_{tst}$	σ_{tst}	**t-test**
WM	65	57605	2841	+	57934	4733	+
WM+T	65	18602	1211	+	22666	3386	+
WM+S	40.8	41086	1322	+	59942	4931	+
WM+TS	41.9	14987	391	+	18973	3772	+
NSGAII	41.0	14488	965	+	18419	3054	+
NSGAII$_{ACC}$	48.1	16321	1636	+	20423	3138	+
SPEA2	**33**	13272	1265	+	17533	3226	+
SPEA2$_{ACC}$	34.5	**11081**	1186	*	**14161**	2191	*

with several single objective based methods. This way to work differs from the previous works in the specialized literature, in which one or several Pareto fronts are presented and an expert should then select one solution. Our main aim following this approach is to compare the same algorithm by only considering an accuracy objective (WM+TS) with the most accurate solution found by the multi-objective ones in order to see if the Pareto fronts obtained are not only wide but also optimal (similar solutions to that obtained by WM+TS should be included in the final Pareto).

5.2. *Results and analysis*

The results obtained by the analyzed methods are shown in Table 3, where $\#R$ stands for the number of rules, MSE$_{tra}$ and MSE$_{tst}$ respectively for the average error obtained over the training and test data, σ for the standard deviation and *t-test* for the results of applying a *test t-student* (with 95 percent confidence) in order to ascertain whether differences in the performance of the proposed approach are significant when compared with that of the other algorithms in the table. The interpretation of this column is:

$\star$ represents the best average result.
$+$ means that the best result has better performance than that of the corresponding row.

Analysing the results showed in Table 3 we can highlight the following facts:

- SPEA2$_{ACC}$ gets an important reduction of the mean square error respect to that obtained by the classic methods and NSGA-II. Furthermore, this algorithm improves the results obtained by SPEA2 with only 1.5 more rules.
- The models obtained by SPEA2$_{ACC}$ seem to show very good trade-off between interpretability and accuracy.

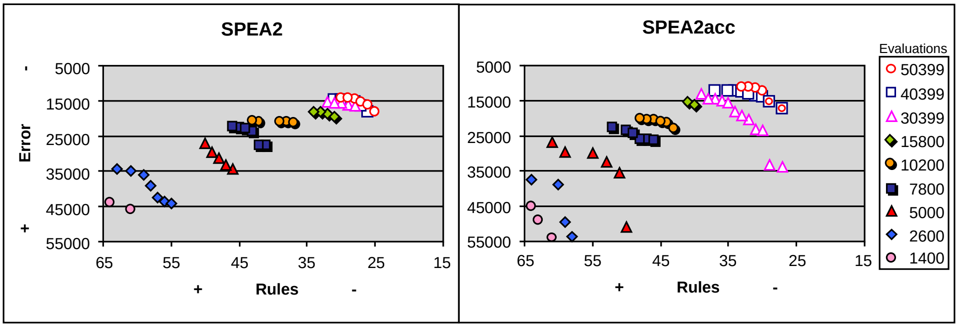


Fig. 6. Pareto fronts of $SPEA2$ and $SPEA2_{acc}$.

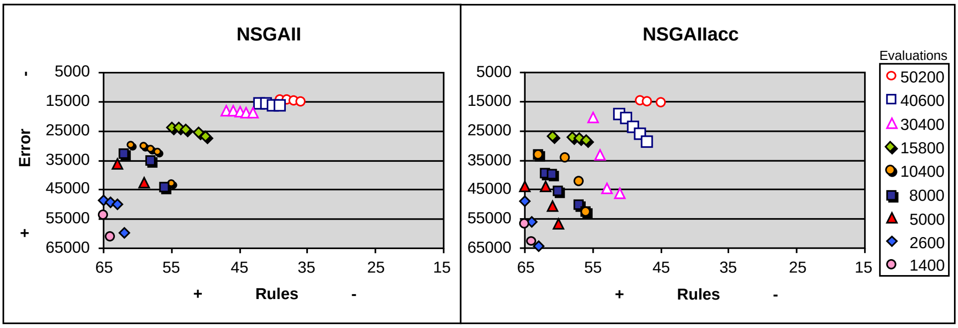


Fig. 7. Pareto fronts of $NSGAII$ and $NSGAII_{acc}$.

- NSGAII and NSGAII$_{ACC}$ present a not so good performance in this particular problem because of the crowding operator which makes very difficult to concentrate the search in the desired Pareto zone.

Moreover, notice the large search space that involves this problem. There are some initial rules that should be removed since they do not cooperate in a good way with the remaining ones. Even in the case of only considering an accuracy-based objective (WM+TS), the large search space that supposes the tuning of parameters makes very difficult to remove these kinds of rules since bad rules are tuned together with the remaining ones searching for their best cooperation. The use of a multi-objective approach favours a better selection of the ideal number of rules, preserving some rule configurations until the rule parameters are evolved to dominate solutions including bad rules.

In Figures 6 and 7, we can see the Pareto evolution for each algorithm. In figure 6, we can observe that SPEA2$_{ACC}$ mainly explores in the mid part of the evolution (before applying the restarting operator) in order to finally focusing on a specific zone of the Pareto. After restarting, the Pareto is extended in order to continue concentrating the search on the Pareto zone presenting solutions with less number of rules but still accurate.

In the remaining methods, Figures 6 and 7, we can see as the Pareto moves along without having a big extension, which does not allow to obtain very good results even in the case of NSGA-II.

6. Conclusions

Taking into account the results showed in the previous section, we can conclude that the models obtained by the proposed method present a better trade-off between interpretability and accuracy than the remaining ones. By searching for a good configuration of rules (only removing rules with little importance) and by tuning the parameter for a small set of rules, the proposed algorithm has obtained models even with a better accuracy than those obtained by methods only guided by measures of accuracy. In this way, the results obtained have shown that the use of MOEAs can represent a way to obtain even more accurate and simpler linguistic models than those obtained by only considering performance measures.

On the other hand, the proposed algorithm (SPEA2_{ACC}) could be of interest in problems that, although presenting a multi-objective nature, need as solution not all the Pareto frontier but only a specific area of it.

References

1. J. Casillas, O. Cordón, F. Herrera and L. Magdalena (Eds.), Accuracy improvements in linguistic fuzzy modelling, *Studies in Fuzziness and Soft Computing* **129** (Springer, Heidelberg, Germany, 2003).
2. R. Alcalá, J. Alcalá-Fdez, J. Casillas, O. Cordón and F. Herrera, Hybrid learning models to get the interpretability-accuracy trade-off in fuzzy modeling, *Soft Computing* **10:9** (2006) 717–734.
3. R. Alcalá, J. Alcalá-Fdez and F. Herrera, A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection, *IEEE Transactions on Fuzzy Systems* (2007), in press.
4. R. Alcalá, J. Alcalá-Fdez, M. J. Gacto and F. Herrera, Rule base reduction and genetic tuning of fuzzy systems based on the linguistic 3-tuples representation, *Soft Computing* **11:5** (2007) 401–419.
5. J. Casillas, O. Cordón, M. J. del Jesus and F. Herrera, Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction, *IEEE Trans. on Fuzzy Systems* **13:1** (2005) 13–29.
6. O. Cordón and F. Herrera, A three-stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples, *International Journal of Approximate Reasoning* **17:4** (1997) 369–407.
7. H. B. Gürocak, A genetic-algorithm-based method for tuning fuzzy-logic controllers, *Fuzzy Sets and Systems* **108:1** (1999) 39–47.
8. Y. Jin, W. von Seelen and B. Sendhoff, On generating FC^3 fuzzy rule systems from data using evolution strategies, *IEEE Transactions on Systems, Man, and Cybernetics — Part B: Cybernetics* **29:4** (1999) 829–845.
9. B. D. Liu, C. Y. Chen and J. Y. Tsao, Design of adaptive fuzzy logic controller based on linguistic-hedge concepts and genetic algorithms, *IEEE Transactions on Systems, Man, and Cybernetics — Part B: Cybernetics* **31:1** (2001) 32–53.

10. D. Nauck and R. Kruse, Neuro-fuzzy systems for function approximation, *Fuzzy Sets and Systems* **101:2** (1999) 261–271.
11. C. A. Coello, D. A. Van Veldhuizen and G. B. Lamont, *Evolutionary algorithms for solving multi-objective problems* (Kluwer Academic Publishers, New York, 2002).
12. C. Coello, Recent trends in evolutionary multiobjective optimization, in A. Abraham, L. Jain, R. Goldberg (Ed.), *Evolutionary Multiobjective Optimization: Theoretical Advances and Applications* (Springer-Verlag, London, 2005) 7–32.
13. C. Coello, G. Toscano and E. Mezura, Current and future research trends in evolutionary multiobjective optimization, in: M. Graña, R. Duro, A. d'Anjou and P. P. Wang (Eds.), *Information Processing with Evolutionary Algorithms: From Industrial Applications to Academic Speculations* (Springer-Verlag, London, 2005) 213–231.
14. C. Coello, Evolutionary multiobjective optimization: A historical view of the field, *IEEE Computational Intelligence Magazine* **1:1** (2006) 28–36.
15. E. Zitzler and L. Thiele, Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach, *IEEE Transactions on Evolutionary Computation* **3:4** (1999) 257–271.
16. O. Cordon, F. Herrera, M. J. del Jesus and P. Villar, A multiobjective genetic algorithm for feature selection and granularity learning in fuzzy-rule based classification systems, *Proceedings of IX IFSA World Congress and XX NAFIPS International Conference* (Vancouver, Canada, 2001) 1253–1258.
17. H. Ishibuchi, T. Murata and I. B. Turksen, Selecting linguistic classification rules by two-objective genetic algorithms, *Proceedings of IEEE-SMC'95* (Vancouver, Canada, 1995) 1410–1415.
18. H. Ishibuchi and T. Murata, Minimizing the fuzzy rule base and its performance by a multi-objective genetic algorithm, *Proceedings of 6th FUZZ-IEEE* (Barcelona, Spain, 1997) 259–264.
19. H. Ishibuchi, T. Murata and M. Gen, Performance evaluation of fuzzy rule-based classification systems obtained by multi-objective genetic algorithms, *Computers & Industrial Engineering* **35:3-4** (1998) 575–578.
20. H. Ishibuchi, T. Nakashima and T. Murata, Three-objective genetics-based machine learning for linguistic rule extraction, *Information Sciences* **136** (2001) 109–133.
21. H. Ishibuchi, T. Nakashima and T. Murata, Multiobjective Optimization in Linguistic Rule Extraction from Numerical, *Proceedings of First International Conference on Evolutionary Multi-Criterion Optimization, Lecture Notes in Computer Science* **1993** (Zurich, Switzerland, 2001) 588–602.
22. H. Ishibuchi and T. Yamamoto, Interpretability Issues in Fuzzy Genetics-Based Machine Learning for Linguistic Modelling, in: J. Lawry, J. G. Shanahan and A. L. Ralescu (Eds.), *Modelling with Words: Learning, Fusion, and Reasoning within a Formal Linguistic Representation Framework, Lecture Notes in Computer Science 2873* (Springer-Verlag, Berlin, Heidelberg, 2003) 209–228.
23. H. Ishibuchi and T. Yamamoto, Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining, *Fuzzy Sets and Systems* **141:1** (2004) 59–88.
24. H. Ishibuchi and Y. Nojima, Accuracy-complexity tradeoff analysis by multiobjective rule selection, *Proceedings of ICDM 2005 Workshop on Computational Intelligence in Data Mining* (Houston, Texas, 2005) 39–48.
25. K. Narukawa, Y. Nojima and H. Ishibuchi, Modification of evolutionary multiobjective optimization algorithms for multiobjective design of fuzzy rule-based classification systems, *Proceedings of 2005 IEEE International Conference on Fuzzy Systems* (Reno, USA, 2005) 809–814.

26. E. Zitzler, M. Laumanns and L. Thiele, SPEA2: Improving the strength pareto evolutionary algorithm for multiObjetive optimization, *Proc. Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems*, Barcelona, Spain, 2001, 95–100.
27. K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Trans. on Evolutionary Computation* **6:2** (2002) 182–197.
28. E. H. Mamdani, Applications of fuzzy algorithms for control a simple dynamic plant, *Proceedings of the IEEE* **121:12** (1974) 1585–1588.
29. E. H. Mamdani and S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, *International Journal of Man-Machine Studies* **7** (1975) 1–13.
30. T. Takagi and M. Sugeno, Fuzzy identification of systems and its application to modeling and control, *IEEE Transactions on Systems, Man, and Cybernetics* **15** (1985) 116–132.
31. J. Casillas, O. Cordón, F. Herrera and L. Magdalena (Eds.), *Interpretability issues in fuzzy modeling*, Studies in Fuzziness and Soft Computing **128** (Springer, Heidelberg, Germany, 2003).
32. Y. Shi, R. Eberhart and Y. Chen, Implementation of evolutionary fuzzy systems, *IEEE Transactions on Fuzzy Systems* **7:2** (1999) 109–119.
33. R. Alcalá, J. Alcalá-Fdez, F. Herrera and J. Otero, Genetic learning of accurate and compact fuzzy rule based systems based on the 2-tuples linguistic representation, *International Journal of Approximate Reasoning* **44:1** (2007) 45–64.
34. W. Pedrycz, R. R. Gudwin and F. A. C. Gomide, Nonlinear context adaptation in the calibration of fuzzy sets, *Fuzzy Sets and Systems* **88:1** (1997) 91–97.
35. J. Espinosa and J. Vandewalle, Constructing fuzzy models with linguistic integrity from numerical data-AFRELI algorithm, *IEEE Transactions on Fuzzy Systems* **8:5** (2000) 591–600.
36. O. Cordón, M. J. del Jesus and F. Herrera, Genetic learning of fuzzy rule-based classification systems cooperating with fuzzy reasoning methods, *International Journal of Intelligent Systems* **13** (1998) 1025–1053.
37. A. González and R. Pérez, A study about the inclusion of linguistic hedges in a fuzzy rule learning algorithm, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **7:3** (1999) 257–266.
38. R. Alcalá, J. R. Cano, O. Cordón, F. Herrera, P. Villar and I. Zwir, Linguistic modeling with hierarchical systems of weighted linguistic rules, *International Journal of Approximate Reasoning* **32:2-3** (2003) 187–215.
39. J. S. Cho and D. J. Park, Novel fuzzy logic control based on weighting of partially inconsistent rules using neural network, *Journal of Intelligent and Fuzzy Systems* **8** (2000) 99–110.
40. N. R. Pal and K. Pal, Handling of inconsistent rules with an extended model of fuzzy reasoning, *Journal of Intelligent and Fuzzy Systems* **7** (1999) 55–73.
41. T. P. Hong and C. Y. Lee, Effect of merging order on performance of fuzzy induction, *Intelligent Data Analysis* **3:2** (1999) 139–151.
42. H. Ishibuchi, K. Nozaki, N. Yamamoto and H. Tanaka, Selecting fuzzy if-then rules for classification problems using genetic algorithms, *IEEE Transactions on Fuzzy Systems* **3:3** (1995) 260–270.
43. T. P. Hong and J. B. Chen, Finding relevant attributes and membership functions, *Fuzzy Sets and Systems* **103:3** (1999) 389–404.
44. H. M. Lee, C. M. Chen, J. M. Chen and Y. L. Jou, An efficient fuzzy classifier with feature selection based on fuzzy entropy, *IEEE Transactions on Systems, Man, and*

Cybernetics — Part B: Cybernetics **31:3** (2001) 426–432.

45. L. J. Eshelman and J. D. Schaffer, Real-coded genetic algorithms and interval-schemata, *Foundations of Genetic Algorithms* **2** (1993) 187–202.
46. L. J. Eshelman, The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination, *Foundations of Genetic Algorithms* **1** (1991) 265–283.
47. L. X. Wang and J. M. Mendel, Generating fuzzy rules by learning from examples, *IEEE Trans. on Systems, Man, and Cybernetics* **22:6** (1992) 1414–1427.
48. O. Cordón, F. Herrera and L. Sánchez, Computing the Spanish Medium Electrical Line Maintenance Costs by means of Evolution-Based Learning Processes, *Proc. International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems*, Castellón, Spain, 1998, 478–486.
49. O. Cordón, F. Herrera and L. Sánchez, Solving electrical distribution problems using hybrid evolutionary data analysis techniques, *Applied Intelligence* **10** (1999) 5–24.

3.2. *Propuesta de un Algoritmo Evolutivo Multi-Objetivo Avanzado para Problemas de Regresión: Estudio sobre Distintas Alternativas*

- M.J. Gacto, R. Alcalá, F. Herrera, Adaptation and Application of Multi-Objective Evolutionary Algorithms for Rule Reduction and Parameter Tuning of Fuzzy Rule-Based Systems. Soft Computing 13:5 (2009) 419-436, doi:10.1007/s00500-008-0359-z.

    - Estado: Publicado.
    - Índice de Impacto (JCR 2009): 1,328.
    - Área de Conocimiento: Computer Science, Artificial Intelligence. Ranking 51 / 102.
    - Área de Conocimiento: Computer Science, Interdisciplinary Applications. Ranking 41 / 95.
    - Citas: 11.

Soft Comput (2009) 13:419–436
DOI 10.1007/s00500-008-0359-z

FOCUS

Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems

María José Gacto · Rafael Alcalá · Francisco Herrera

Published online: 26 August 2008

Abstract Recently, multi-objective evolutionary algorithms have been applied to improve the difficult tradeoff between interpretability and accuracy of fuzzy rule-based systems. It is known that both requirements are usually contradictory, however, these kinds of algorithms can obtain a set of solutions with different trade-offs. This contribution analyzes different application alternatives in order to attain the desired accuracy/interpr-etability balance by maintaining the improved accuracy that a tuning of membership functions could give but trying to obtain more compact models. In this way, we propose the use of multi-objective evolutionary algorithms as a tool to get almost one improved solution with respect to a classic single objective approach (a solution that could dominate the one obtained by such algorithm in terms of the system error and number of rules). To do that, this work presents and analyzes the application of six different multi-objective evolutionary algorithms to obtain simpler and still accurate linguistic fuzzy models by performing rule selection and a tuning of the membership functions. The results on two different scenarios show that the use of expert knowledge in the algorithm design process significantly improves the search ability of these algorithms and that they are able to improve both objectives together, obtaining more accurate and at the same time simpler models with respect to the single objective based approach.

M. J. Gacto (✉) · R. Alcalá · F. Herrera
Department of Computer Science and A.I,
University of Granada, Granada 18071, Spain
e-mail: mjgacto@ugr.es

R. Alcalá
e-mail: alcala@decsai.ugr.es

F. Herrera
e-mail: herrera@decsai.ugr.es

1 Introduction

Many automatic techniques have been proposed in the literature to extract a proper set of fuzzy rules from numerical data. Most of these techniques usually try to improve the performance associated to the prediction error without pay a special attention to the system interpretability, an essential aspect of fuzzy rule-based systems (FRBSs). In the last years, the problem of finding the right trade-off between interpretability and accuracy, in spite of the original nature of fuzzy logic, has arisen a growing interest in methods which take both aspects into account (Casillas et al. 2003a, b). Of course, the ideal thing would be to satisfy both criteria to a high degree, but since they are contradictory issues generally it is not possible. A way to do that, is to improve the system accuracy but trying to maintain the interpretability to an acceptable level (Casillas et al. 2003b).

A widely-used technique to improve the accuracy of linguistic FRBSs is the *tuning* of membership functions (MFs) (Alcalá et al. 2006, 2007b, c; Casillas et al. 2003b, 2005), which refines a previous definition of the data base once the rule base has been obtained. The classic approach to perform tuning consists of refining the three definition parameters that identify triangular MFs associated to the labels comprising the initial data base. Although tuning is one of the most powerful techniques to improve the system performance (Casillas et al. 2003b, 2005), sometimes an excessive number of rules is initially considered to reach the highest degree of accuracy. In order to maintain the interpretability to an acceptable level, a recent work (Casillas et al. 2005) has considered the selection of rules together with the tuning of MFs within the same process (not in different stages) and considering performance criteria. In this way, rules are extracted only if it is possible to maintain or even improve the system accuracy. A very interesting conclusion from (Casillas et al. 2005) is that both techniques can present a positive synergy in most of the cases (similar or more accurate models could be obtained by reducing the number of rules) when they are combined within the same process.

On the other hand, since this approach presents a multi-objective nature the use of multi-objective evolutionary algorithms (MOEAs) (Coello et al. 2002; Deb 2001) to obtain a set of solutions with different degrees of accuracy and number of rules could represent an interesting way to work (by considering both characteristics as objectives). Although there are some works in the literature using standard or specific MOEAs to improve the difficult trade-off between interpretability and accuracy of Mamdani FRBSs (Cococcioni et al. 2007; Cordon et al. 2001; Ishibuchi et al. 1997, 2001; Ishibuchi and Yamamoto 2003, 2004; Narukawa et al. 2005), practically all these works were applied to classification problems trying to obtain the complete Pareto (set of non-dominated solutions with different trade-off) by selecting or learning the set of rules better representing the example data, i.e., improving the system classification ability and decreasing the system complexity but not considering learning (Alcalá et al. 2007a) or tuning (Alcalá et al. 2006, 2007a, b; Casillas et al. 2003a, 2005) of the fuzzy system parameters, which involves another degree of trade-off and type of Pareto front, a more complicated search space and therefore needs different considerations with respect to the works in the existing literature.

Indeed, to directly apply the most recognized MOEAs for general use in order to perform together tuning and rule selection could present some important problems. As David Goldberg stated in (Goldberg 2000), the integration of single methods into hybrid intelligent systems goes beyond simple combinations. For him, the future of Computational Intelligence "lies in the careful integration of the best constituent technologies", and subtle integration of the abstraction power of fuzzy systems and the innovative power of genetic systems requires a design sophistication that goes further than putting everything together. This is the case when parameter tuning and rule selection are performed by directly applying the most known MOEAs for general use, where several problems arise due to the complex search space concerning this problem.

The main problem is that it is practically impossible to obtain the complete optimal Pareto. This is due to several reasons:

1. There exist a lot of different subsets of rules with more or less the same number of rules (different rule configurations) but representing really different or alternative tuning possibilities.
2. It is easier to decrease the number of rules than to reduce the system error (which is more dependent of the tuning task). This provokes a faster tuning of the simplest solutions before exploring more promising rule configurations (which are dominated by such premature solutions).
3. The obtained parameters (in general) tends to be optimal for these premature solutions making difficult the appearance of better alternative solutions.

In this way, it is necessary to include any expert knowledge in the MOEA design process. An adequate application of standard MOEAs could partially deal with this problem by focusing the search in the most interesting zone of the Pareto frontier. Taking into account that non-dominated solutions with a small number of rules and high errors are not interesting since they have not the desired trade-off between accuracy and interpretability, we could focus the search only in the Pareto zone with the most accurate solutions trying to obtain the least possible number of rules. The best way to do this is to start with solutions selecting all the possible rules, which favours a progressive extraction of bad rules (those that do not improve with the tuning of parameters), only by means of the mutation at the beginning and then by means of the crossover.

A secondary problem is that it is difficult to obtain very accurate solutions by favoring the crossing of solutions with very different rule configurations (those in the Pareto), which should obtain the best accuracy by learning different parameters for the MFs. Although this is not the major problem (MOEAs can obtain good results considering the established mechanisms), significant improvements could be achieved by addressing this problem in the proper way, i.e., standard algorithms can be specifically improved to perform rule selection and tuning together. A way to do that is to establish mating restrictions. However, again it should be done based on the experience by taking into account that exploration and exploitation are both mainly needed at different stages. In this way, a new method was recently proposed in (Alcalá et al. 2007d), which by modifying the Strength Pareto Evolutionary Algorithm 2 (SPEA2) (Zitzler et al. 2001) progressively concentrates the search in the most promising solutions, allowing exploration at first and favoring the exploitation of the most accurate solutions at the end (the Accuracy-Oriented SPEA2, $SPEA2_{Acc}$). Another possibility could be the application of two versions of the well-known Nondominated Sorting Genetic Algorithm II (NSGA-II) (Deb et al.

2002) for finding knees (Branke et al. 2004) (theoretically the most promising Pareto zones in these kinds of problems), since the modifications proposed in (Alcalá et al. 2007d) were not successful by considering the NSGA-II approach.

Our main aim in this contribution is to analyze different alternatives in order to attain the desired accuracy/interpretability balance by maintaining the impro-ved accuracy that a tuning of MFs could give but trying to obtain more compact models. In this way, we propose the use of MOEAs as the tool to get almost one improved solution with respect to the classic single objective algorithm (a solution that could dominate the one obtained by such algorithm in terms of the system error and the number of rules). To do that, this work presents and analyzes the application of six different MOEAs to obtain simpler and still accurate linguistic fuzzy models by performing rule selection and a classic tuning of the MF parameters (an analysis on the classic tuning could help to extend the better approaches in order to consider other kinds of techniques or new interpretability measures for further works, e.g., another tuning types, learning, etc.). These methods are the well-known SPEA2, NSGA-II and two versions of NSGA-II for finding knees (standard MOEAs adapting and applying proper genetic operators), and two extended MOEAs for specific application, $SPEA2_{Acc}$ and an extension of it proposed in this paper that by applying a more intelligent crossover operator (specific for this problem) is able to extract more useful information from parents with different configurations, $SPEA2_{Acc^2}$. The results on two different scenarios show that the use of expert knowledge in the MOEAs design process significantly improves the search ability of these algorithms.

In order to do that, this contribution is arranged as follows. Next section presents a brief study of the existing MOEAs for general purpose which usually are modified or directly applied to obtain FRBSs with good interpretability-accuracy trade-off. In order to show the main differences with the previous works, Sect. 3 briefly analyzes the state of the art on the use of MOEAs to get the desired trade-off in different application areas of FRBSs. In Sect. 4, we describe the different MOEAs and appropriate genetic operators for their proper application. Section 5 shows an ex-pe-ri-men-tal study of these me-thods in two complex and interesting problems. Finally, Sect. 6 points out some conclusions and further research lines.

2 Multi-objective evolutionary algorithms

Evolutionary algorithms simultaneously deal with a set of possible solutions (the so-called population) which allows to find several members of the Pareto optimal set in a single run of the algorithm. Additionally, they are not too susceptible to the shape or continuity of the Pareto front (e.g., they can easily deal with discontinuous and concave Pareto fronts).

The first hint regarding the possibility of using evolutionary algorithms to solve a multi-objective problem appears in a PhD thesis from 1967 (Rosenberg 1967) in which, however, no actual MOEA was developed (the multi-objective problem was restated as a single-objective problem and solved with a genetic algorithm). David Schaffer is normally considered to be the first to have designed an MOEA during the mid-1980s (Schaffer 1985). Schaffer's approach, called vector evaluated genetic algorithm (VEGA) consists of a simple genetic algorithm with a modified selection mechanism. However, VEGA had a number of problems from which the main one had to do with its inability to retain solutions with acceptable performance, perhaps above average, but not outstanding for any of the objective functions.

After VEGA, the researchers designed a first generation of MOEAs characterized by its simplicity where the main lesson learned was that successful MOEAs had to combine a good mechanism to select non-dominated individuals (perhaps, but not necessarily, based on the concept of Pareto optimality) combined with a good mechanism to maintain diversity (fitness sharing was a choice, but not the only one). The most representative MOEAs of this generation are the following: Nondominated Sorting Genetic Algorithm (NSGA) (Srinivas and Deb 1994), Niched-Pareto Genetic Algorithm (NPGA) (Horn et al. 1994) and multi-objective genetic algorithm (MOGA) (Fonseca and Fleming 1993).

A second generation of MOEAs started when elitism became a standard mechanism. In fact, the use of elitism is a theoretical requirement in order to guarantee convergence of an MOEA. Many MOEAs have been proposed during the second generation (which we are still living today). However, most researchers will agree that few of these approaches have been adopted as a reference or have been used by others. In this way, SPEA2 and NSGA-II can be considered as the most representative MOEAs of the second generation, also being of interest some others as the Pareto Archived Evolution Strategy (PAES) (Knowles and Corne 2000). Table 1 shows a summary of the most representative MOEAs of both generations.

Finally, we have to point out that nowadays NSGA-II is the paradigm within the MOEA research community since the powerful crowding operator that this algorithm uses usually allows to obtain the widest Pareto sets in a great variety of problems, which is a very appreciated property in this framework. In this way, the question is: "Is NSGA-II the best MOEA to get the desired

Table 1 Classification of MOEAs

Reference	MOEA	1st Gen.	2nd Gen.
Fonseca and Fleming (1993)	MOGA	✔	
Horn et al. (1994)	NPGA	✔	
Srinivas and Deb (1994)	NSGA	✔	
Coello and Toscano (1993)	micro-GA		✔
Erickson et al. (1993)	NPGA 2		✔
Deb et al. (2002)	NSGA-II		✔
Knowles and Corne (2000)	PAES		✔
Corne et al. (2000, 2001)	PESA & PESA-II		✔
Zitzler and Thiele (1999), Zitzler et al. (2001)	SPEA & SPEA2		✔

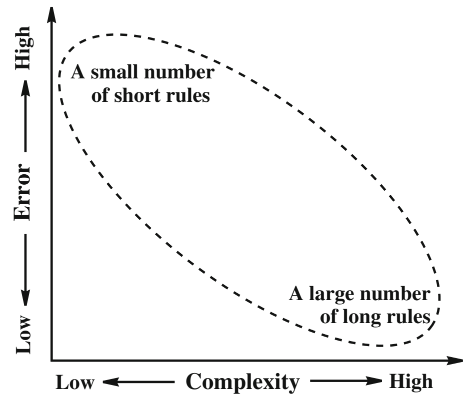


Fig. 1 Trade-off between the error and the interpretability of rule sets

interpretability-accuracy trade-off of FRBSs following our concrete approach?" (tuning and rule selection). In this work, we analyze the behavior of this algorithm, SPEA2 and two versions of NSGA-II developed to find knees (Branke et al. 2004) in the optimal Pareto front, which could be a better way to find still accurate solutions but presenting the least possible number of rules. Additionally, we consider two algorithms based on SPEA2 that are specifically designed to address our problem. Next section presents the state-of-the-art on the use of the MOEAs to get this difficult trade-off in order to see how different researchers have faced this problem.

3 Use of MOEAs to get the interpretability: accuracy trade-off of FRBSs

As mentioned, MOEAs generate a family of equally valid solutions, where each solution tends to satisfy a criterion to a higher extent than another. For this reason, MOEAs have been also applied to improve the difficult trade-off between interpretability and accuracy of FRBSs, where each solution in the Pareto front represents a different trade-off between interpretability and accuracy (see Fig. 1).

The most continuous and prolific research activity in the application of MOEAs to Mamdani FRBS generation for finding the accuracy-interpretability trade off has been certainly performed by Ishibuchi's group. Earlier works (Ishibuchi et al. 1997) were devoted to the application of simple MOEAs of the first generation to perform a rule selection on an initial set of classification rules involving "do not care" conditions and considering two different objectives (classification accuracy and number of rules). Then, a third objective was also included in order to minimize the length of the rules by rule selection (Ishibuchi et al. 2001) or rule learning (Ishibuchi et al. 2001). In (Ishibuchi and Yamamoto 2004), they apply a better MOEA, the Multi-Objective Genetic Local Search (Ishibuchi and Murata 1996) (MOGLS), following the same approach for rule selection with three objectives. And finally, two algorithms based on an MOEA of the second generation (NSGA-II) have been proposed respectively for rule selection (Narukawa et al. 2005) and rule learning (Ishibuchi and Nojima 2007) considering the same concepts. In the literature, we can also find some papers of other researchers in this topic. For instance in (Cordon et al. 2001), Cordon et al. use MOGA for jointly performing feature selection and fuzzy set granularity learning with only two objectives.

At this point, we can see that all the methods mentioned were applied to classification problems for rule selection or rule learning, without learning or tuning the MFs that were initially fixed. Most of the works in this topic only consider quantitative measures of the system complexity in order to improve the interpretability of such systems, rarely considering qualitative measures. Moreover, MOEAs considered were slight modifications of MOEAs proposed for general use (MOGA, NSGA-II, etc.) or based on them. Notice that, although NSGA-II improves the results with respect to other MOEAs, since to cross non-dominated rule sets with very different numbers of rules and different rule structures (forced by NSGA-II crowding operator) usually gives a bad accuracy, this MOEA could need of an adaptation to favor the cross of similar solutions in order to also get good results for the accuracy objective (see Narukawa et al. 2005). A possibility could be the use of similarity measures as the work in Narukawa et al. (2005) (by also favoring the crossover of similar solutions), and other possibility could be to modify the crowding measure as the work in Branke et al. (2004) to find knees in multi-objective optimization problems.

On the other hand, there are a few works in the framework of fuzzy modeling for regression problems. In

Ishibuchi and Yamamoto (2003), authors show how a simple MOGA can be applied to a three-objective optimization problem to obtain Mamdani FRBSs. In Cococcioni et al. (2007), an adaptation of the efficient (2 + 2)PAES (Knowles and Corne 2000) has been applied to the identification of Mamdani FRBSs for regression problems by considering two minimization objectives (the system error and the number of variables involved in the antecedent of the obtained rules). Again, these approaches do not consider learning or tuning of parameters. However, a new method was recently proposed in Alcalá et al. (2007d) to perform rule selection and parameter tuning of Mamdani FRBSs by establishing mating restrictions to concentrate the search in the most promising solutions, allowing exploration at first and favoring the exploitation of the most accurate solutions at the end ($SPEA2_{Acc}$). This last approach will be also analyzed and described in this contribution.

Some applications of MOEAs have been also discussed in the literature to improve the difficult trade-off between accuracy and interpretability of Takagi–Sugeno models (Takagi and Sugeno 1985). In Jimenez et al. (1993), Wang et al. (2005a, b), accuracy, interpretability and compactness have been considered as objectives to obtain interpretable and very accurate Takagi–Sugeno models. However, since Takagi–Sugeno models have a linear function in the consequent part of each fuzzy rule, they are close to accuracy representing another type of trade-off with less interpretable models (Ishibuchi and Yamamoto 2003). For this reason, the type of rule most used to achieve the trade-off between accuracy and complexity are the fuzzy rules with linguistic terms in both the antecedent and consequent parts, i.e., Mamdani rules (Mamdani and Assilian 1975).

4 Six different MOEAs for rule selection and tuning of membership functions

As we explain in the previous section most works in the field of fuzzy systems are applied to classification problems by learning or selecting rules, not considering tuning of the MF parameters. The main reason of this fact is that a tuning of parameters implies a lost of the interpretability to some degree. However, it is known that this way to work greatly improves the performance of the linguistic models so obtained, being another alternative to improve the interpretability-accuracy trade-off. For this reason, we would like to show six examples of applications that focus the research in the linguistic fuzzy modeling area, in order to evaluate the performance of MOEAs in a field which is still less explored, and with the objective of inject some ideas or recommendations in this open topic (improvement of the interpretability of very accurate models).

The proposed algorithms will perform rule selection from a given fuzzy rule set together with a parametric tuning of the MFs. To do that, we apply the most used multi-objective algorithms of the second generation, SPEA2 (Zitzler et al. 2001) and NSGA-II (Deb et al. 2002), and two versions of NSGA-II (Branke et al. 2004) for finding knees. Moreover we consider two extended MOEAs for specific application to this concrete problem, $SPEA2_{Acc}$ in Alcalá et al. (2007d), and an extension of that, called $SPEA2_{Acc^2}$. All of them consider two different objectives, system error and number of rules.

In the next subsections, we present SPEA2, NSGA-II, NSGA-II_A, NSGA-II_U and $SPEA2_{Acc}$ algorithms and we propose $SPEA2_{Acc^2}$ applied for linguistic fuzzy modeling. At first, the common components of these algorithms are proposed and then the main steps and characteristic of them are described.

4.1 Main components of the algorithms

As mentioned, we propose six algorithms to perform rule selection and tuning of MFs and with the aim of improving the desired trade-off between interpretability and accuracy. In the following, the common components needed to apply these algorithms in this concrete problem are explained. They are coding scheme, initial gene pool, objectives and genetic operators:

- Coding scheme and initial gene pool

A double coding scheme for both *rule selection* (C_S) and *tuning* (C_T) is used:

$C^p = C_S^p C_T^p$

In the $C_S^p = (c_{S1}, ..., c_{Sm})$ part, the coding scheme consists of binary-coded strings with size m (with m being the number of initial rules). Depending on whether a rule is selected or not, values '1' or '0' are respectively assigned to the corresponding gene. In the C_T part, a real coding is considered, being m^i the number of labels of each of the n variables comprising the data base,

$$C_i = (a_1^i, b_1^i, c_1^i, \ldots, a_{m^i}^i, b_{m^i}^i, c_{m^i}^i), \quad i = 1, \ldots, n,$$
$$C_T^p = C_1 C_2, \ldots, C_n.$$

The initial population is obtained with all individuals having all genes with value '1' in the C_S part. And in the C_T part the initial data base is included as first individual. The remaining individuals are generated at random within the corresponding variation intervals. Such intervals are calculated from the initial data base. For each MF,

$C_i^j = (a^j, b^j, c^j)$, the variation intervals are calculated in the following way:

$$[I_{a^j}^l, I_{a^j}^r] = [a^j - (b^j - a^j)/2, a^j + (b^j - a^j)/2]$$
$$[I_{b^j}^l, I_{b^j}^r] = [b^j - (b^j - a^j)/2, b^j + (c^j - b^j)/2]$$
$$[I_{c^j}^l, I_{c^j}^r] = [c^j - (c^j - b^j)/2, c^j + (c^j - b^j)/2]$$

- Objectives

Two objectives are minimized for this problem: the number of rules (interpretability) and the mean squared error (accuracy),

$$\mathrm{MSE} = \frac{1}{2 \cdot |E|} \sum_{l=1}^{|E|} (F(x^l) - y^l)^2,$$

with $|E|$ being the size of a data set E, $F(x^l)$ being the output obtained from the FRBS decoded from the said chromosome when the l-th example is considered and y^l being the known desired output. The fuzzy inference system considered to obtain $F(x^l)$ is the *center of gravity weighted by the matching* strategy as defuzzification operator and the *minimum t-norm* as implication and conjunctive operators.

- Genetic operators

The crossover operator depends on the chromosome part where it is applied: the BLX-0.5 (Eshelman and Schaffer 1993) in the C_T part and the HUX (Eshelman 1991) in the C_S part.

Finally, four offspring are generated by combining the two from the C_S part with the two from the C_T part (the two best replace to their parent). The mutation operator changes a gene value at random in the C_S and C_T parts (one in each part) with probability P_m.

- Importance of the initial population

Besides, we have to highlight that the way to create the solutions of the initial population for the part of rule selection is a very important factor. Usually, a genetic algorithm generates the initial population totally at random (random selection of the initial rules). However, in this case, to get solutions with a high accuracy we should not lose rules that could present a positive cooperation once their MF parameters have been evolved. The best way to do this is to start with solutions selecting all the possible rules which favors a progressive extraction of bad rules (those that do not improve with the tuning of parameters), only by means of the mutation at the beginning and then by means of the crossover. Different proofs were performed considering a completely random initialization, obtaining simpler solutions but with really worse error values in training and test.

4.2 SPEA2 based approach

SPEA2 algorithm (Zitzler et al. 2001) was designed to overcome the problems of its predecessor for general multi-objective optimization, SPEA algorithm (Zitzler and Thiele 1999). In contrast with SPEA, SPEA2: (1) incorporates a fine-grained fitness assignment strategy which takes into account for each individual the number of individuals that it dominates and the number of individuals by which it is dominated; (2) uses the nearest neighbour density estimation technique which guides the search more efficiently; (3) has an enhanced archive truncation method which guarantees the preservation of boundary solutions. Next, we briefly describe the complete SPEA2 algorithm.

SPEA2 uses a fixed population and archive size. The population forms the current base of possible solutions, while the archive contains the current solutions. The archive is constructed and updated by copying all non-dominated individuals in both archive and population into a temporary archive. If the size of this temporary archive differs from the desired archive size, individuals are either removed or added as necessary. Individuals are added by selecting the best dominated individuals, while the removal process uses a heuristic clustering routine in the objective space. The motivation for this is that one would like to try to ensure that the archive contents represent distinct parts of the objective space. Finally, when selecting individuals for participating in the next generation all candidates are selected from the archive using a binary tournament selection scheme.

Considering the components defined and the descriptions of the authors in Zitzler et al. (2001), SPEA2 algorithm consists of the next steps:

$$\begin{aligned}
&\text{Input}: N(\text{population size}),\\
&\overline{N}(\text{external population size}),\\
&T(\text{maximum number of generations}).\\
&\text{Output}: \mathrm{A}(\text{non-dominated set}).
\end{aligned}$$

1. Generate an initial population P_0 and create the empty external population $\overline{P}_0 = \emptyset$.
2. Calculate fitness values of individuals in P_t and $\overline{P}_t$.
3. Copy all non-dominated individuals in $P_t \cup \overline{P}_t$ to $\overline{P}_{t+1}$. If $|\overline{P}_{t+1}| > \overline{N}$ apply truncation operator. If $|\overline{P}_{t+1}| < \overline{N}$ fill with dominated in $P_t \cup \overline{P}_t$.
4. If $t \geq T$, return A and stop.
5. Perform binary tournament selection with replacement on $\overline{P}_{t+1}$ in order to fill the mating pool.
6. Apply recombination (BLX–HUX) and mutation operators to the mating pool and set P_{t+1} to the resulting population. Go to step 2 with $t = t + 1$.

4.3 NSGA-II based approach

NSGA-II algorithm (Deb et al. 2002) is one of the most well-known and frequently-used MOEAs for general multi-objective optimization in the literature. As in other evolutionary algorithms, first NSGA-II generates an initial population. Then an offspring population is generated from the current population by selection, crossover and mutation. The next population is constructed from the current and offspring populations. The generation of an offspring population and the construction of the next population are iterated until a stopping condition is satisfied. NSGA-II algorithm has two features, which make it a high-performance MOEA. One is the fitness evaluation of each solution based on Pareto ranking and a crowding measure, and the other is an elitist generation update procedure.

Each solution in the current population is evaluated in the following manner. First, Rank 1 is assigned to all non-dominated solutions in the current population. All solutions with Rank 1 are tentatively removed from the current population. Next, Rank 2 is assigned to all non-dominated solutions in the reduced current population. All solutions with Rank 2 are tentatively removed from the reduced current population. This procedure is iterated until all solutions are tentatively removed from the current population (i.e., until ranks are assigned to all solutions). As a result, a different rank is assigned to each solution. Solutions with smaller ranks are viewed as being better than those with larger ranks. Among solutions with the same rank, an additional criterion called a crowding measure is taken into account.

The crowding measure for a solution calculates the distance between its adjacent solutions with the same rank in the objective space. Less crowded solutions with larger values of the crowding measure are viewed as being better than more crowded solutions with smaller values of the crowding measure.

A pair of parent solutions are selected from the current population by binary tournament selection based on the Pareto ranking and the crowding measure. When the next population is to be constructed, the current and offspring populations are combined into a merged population. Each solution in the merged population is evaluated in the same manner as in the selection phase of parent solutions using the Pareto ranking and the crowding measure. The next population is constructed by choosing a specified number (i.e., population size) of the best solutions from the merged population. Elitism is implemented in NSGA-II algorithm in this manner.

Considering the components previously defined and the descriptions of the authors in Deb et al. (2002), NSGA-II consists of the next steps:

1. A combined population R_t is formed with the initial parent population P_t and offspring population Q_t (initially empty).
2. Generate all non-dominated fronts $F = (F_1, F_2, ...)$ of R_t.
3. Initialize $P_{t+1} = 0$ and $i = 1$.
4. Repeat until the parent population is filled.
5. Calculate crowding-distance in F_i.
6. Include i-th non-dominated front in the parent population.
7. Check the next front for inclusion.
8. Sort in descending order using crowded-comparison operator.
9. Choose the first $(N-|P_{t+1}|)$ elements of F_i.
10. Use selection, crossover (BLX–HUX) and mutation to create a new population Q_{t+1}.
11. Increment the generation counter.

4.4 NSGA-II with angle-measure based approach

As mentioned, the performance of NSGA-II relies on two measures when comparing individuals: the first is the non-domination rank and, if two individuals have the same non-domination rank, as a secondary criterion, a crowding measure is used.

In Branke et al. (2004), authors presented a different version of NSGA-II in order to find knees in the Pareto front by modifying the secondary criterion, and replacing the crowding measure by either an angle-based measure or an utility-based measure. Again, this algorithm was proposed for multi-objective optimization in general. However, in our case, a knee could represent the best compromise between accuracy and number of rules. So we propose the use of these kinds of measures to search for these interesting Pareto zones in our concrete problem. In this subsection, the use of the angle-based measure is explained in order to replace the crowding measure of NSGA-II.

In the case of only two objectives, the trade-offs in either direction can be estimated by the slopes of the two lines through an individual and its two neighbors. The angle between these slopes can be regarded as an indication of whether the individual is at a knee or not. For an illustration, consider Fig. 2. Clearly, the larger the angle α between the lines, the worse the trade-offs in either direction, and the more clearly the solution can be classified as a knee.

More formally, to calculate the angle measure for a particular individual C^i, we calculate the angle between the individual and its two neighbors, i.e. between (C^{i-1},C^i) and (C^i,C^{i+1}). These three individuals have to be pairwise linearly independent, thus duplicate individuals (individuals

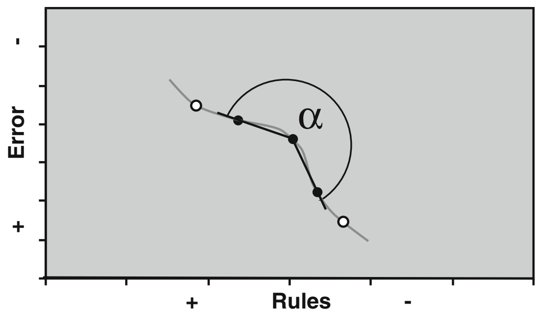


Fig. 2 Calculation of the angle measure

with the same objective function values, which are not prevented in NSGA-II per se) are treated as one and are assigned the same angle-measure. If no neighbor to the left (right) is found, a horizontal (vertical) line is used to calculate the angle. Similar to the standard crowding measure, individuals with a larger angle-measure are preferred.

Calculating the angle measure in 2D is efficient. For more than two objectives, however, it becomes impractical even to just find the neighbors. Thus, we restrict our examination of the angle-based focus to problems with two objectives only. Another important issue is that the values of the different objectives have to be normalized in order to calculate fair angle values. In our case, the sides of the triangles used to compute the final value of α are divided by the difference between the best and the worst values of the corresponding objective in the current Pareto front, as it is done in the original NSGA-II to compute the crowding measure.

4.5 NSGA-II with utility-measure based approach

An alternative measure for a solution's relevance was also proposed in (Branke et al. 2004). This subsection explains the use of this measure (utility-based measure) in order to provide a different way to replace the crowding measure of NSGA-II.

The proposed alternative measure is the expected marginal utility that a solution provides to a decision maker, assuming linear utility functions of the form $U(C, \lambda) = \lambda f_1(C) + (1 - \lambda) f_2(C)$, with all $\lambda \in [0,1]$ being equally likely. For illustration, let us first assume we would know that the decision maker has a particular preference function $U(C, \lambda')$, with some known λ'. Then, we could calculate, for each individual C^i in the population, the decision maker's utility $U(C^i, \lambda')$ of that individual. Clearly, given the choice among all individuals in the population, the decision maker would select the one with the highest utility. Now let us define an individual's marginal utility $U'(C, \lambda')$ as the additional cost the decision maker would have to accept if that particular individual would not be available and he/she would have to settle for the second best, i.e.

$$U'(C^i,\lambda') = \begin{cases} \min_{j\neq i} U(C^j,\lambda') - U(C^i,\lambda') & : i = \arg\min U(C^j,\lambda') \\ 0 & : otherwise \end{cases}$$

The proposed utility measure assumes a distribution of utility functions uniform in the parameter λ in order to calculate the expected marginal utility. For the case of only two objectives, the expected marginal utility can be calculated exactly by integrating over all possible linear utility functions. However, the expected marginal utilities can be approximated simply by sampling, i.e. by calculating the marginal utility for all individuals for a number of randomly chosen utility functions, and taking the average as expected marginal utility. Sampling can be done either randomly or, as was proposed in Branke et al. (2004) in order to reduce variance, in a systematic manner (equidistant values for λ). The number of utility functions used for approximation was called precision of the measure. Authors recommend a precision of at least the number of individuals in the population. Naturally, individuals with the largest overall marginal utility are preferred.

Notice, however, that the assumption of linear utility functions makes it impossible to find knees in concave regions of the non-dominated front. Unlike the angle measure, the utility measure extends easily to more than two objectives, by defining $U(C, \lambda) = \Sigma\lambda_i f_i(C)$ *with* $\Sigma\lambda_i = 1$.

In this paper, the marginal utilities have been computed by sampling, considering equi-distant values for λ and a precision of exactly the number of individuals in the population. As in the case of the angle-measure based approach, the values of the different objectives have to be normalized in order to calculate fair utility values. In our case, objective values considered for computing utility values were normalized considering the best and the worst values of the corresponding objective in the current Pareto front. In this way, an objective value can be normalized as,

$$f_i'(C) = (f_i(C) - f_i^{MIN})/f_i^{MAX},$$

providing values between 0.0 and 1.0.

4.6 Accuracy-oriented based approach: SPEA2$_{Acc}$ algorithm

SPEA2$_{Acc}$ algorithm was very recently proposed in Alcalá et al. (2007d), and is a particularization of SPEA2 based approach presented in Sect. 4.2 to better solve the problem of rule selection and tuning of FRBSs. This algorithm tries to focus the search on the desired Pareto zone, high accuracy with least possible number of rules, proposing two main changes on SPEA2 algorithm with the aim of giving more selective pressure to those solutions that have a high accuracy (crossing dissimilar solutions in principle and similar ones at the end). These changes were also applied and analyzed on NSGA-II in Alcalá et al. (2007d) showing not so good results. The proposed changes are described next:

- A restarting operator is applied exactly at the mid of the algorithm, by maintaining the most accurate individual as the sole individual in the external population ($\overline{P}_{t+1}$ with size 1) and obtaining the remaining individuals in the population (P_{t+1}) with the same rule configuration of the best individual and tuning parameters generated at random within their corresponding variation intervals. This operation is performed in step 4 (see Sect. 4.2) as a second condition, then returning to step 2 with $t = t + 1$. In this way, the search is concentrated only in the desired Pareto zone (similar solutions in a zone with high accuracy).
- In each stage of the algorithm (before and after restarting), the number of solutions in the external population ($\overline{P}_{t+1}$) considered to form the mating pool is progressively reduced, by focusing only on those with the best accuracy. To do that, the solutions are sorted from the best to the worst (considering accuracy as sorting criterion) and the number of solutions considered for selection is reduced progressively from 100% at the beginning to 50% at the end of each stage.

4.7 Extension of SPEA2$_{\text{Acc}}$ algorithm: SPEA2$_{\text{Acc}^2}$

SPEA2$_{\text{Acc}}$ algorithm tries to focus the search in the Pareto zone containing the most accurate solutions. This algorithm represents a good way to obtain more accurate solutions by maintaining only a few more rules with respect to its counterpart (SPEA2). However, sometimes this fact could represent a problem since there are problems in which to obtain accurate solutions could be easy but not so easy to remove unnecessary rules. In this subsection, we propose an extension of this algorithm in order to solve this problem. To do that, we propose two changes based on our experience in this concrete problem:

- An intelligent crossover that is able to profit even more from the corresponding parents, replacing the HUX crossover for the C_S part. To obtain each offspring the following steps are applied:
  1. The BLX crossover is applied to obtain the C_T part of the offspring.
  2. Once the real parameters are obtained determining a the data base, for each gene in the C_S part the corresponding rule is independently extracted from each individual involved in the crossover (offspring and parents 1 and 2). In this way, the same rule is obtained three times with different MFs (those concerning these three individuals).
  3. Euclidean normalized distances are computed between offspring and each parent by only considering the center points (vertex) of the MFs involved in the extracted rules. The differences between each two points are normalized by the amplitude of their respective variation intervals.
  4. The nearest parent is the one that determines if this rule is selected or not for the offspring by directly copying its value in C_S for the corresponding gene.
  5. This process is repeated until all the C_S values are assigned for the offspring.
- Four offspring are obtained repeating this process four times (after considering mutation, only the two most accurate are taken as descendent). By applying this operator, exploration is performed in the C_T part and the C_S part is directly obtained based on the previous knowledge each parent has about the use or not of a specific configuration of MFs for each rule. This avoid to recover a bad rule that was discarded for a concrete configuration of MFs, or allow to recover a good rule that is still considered for a concrete configuration of MFs, increasing the probability of succeed in the selection or elimination of a rule for each concrete configuration of MFs.
- Since a better exploration is performed for the C_S part, the mutation operator does not need to add rules (rules that were eliminated in the parents for a similar bad configuration of the MFs involved in these rules). In this way, once an offspring is generated the mutation operator changes a gene value at random in the C_T part (as in the previous algorithm) and directly sets to zero a gene selected at random in the C_S part (one gene is considered in each part) with probability P_m.

Applying these operators two problems are solved. Firstly, crossing individuals with very different rule configurations is more productive. And secondly, this way to work favors rule extraction since mutation is only engaged to remove unnecessary rules.

5 Experiments

To evaluate the goodness of the proposed approaches, two real-world problems with different complexities (different number of variables and available data) are considered to be solved:

- An *electrical distribution problem* (Cordón et al. 1999) that consists of estimating the maintenance costs of

Table 2 Methods considered for comparison

Method	Ref.	Description
WM	Wang and Mendel (1992)	Wang & Mendel algorithm
T	Casillas et al. (2005)	Tuning of Parameters
S	Casillas et al. (2005)	Rule Selection
TS	Casillas et al. (2005)	Tuning & Selection
Application of standard MOEAs for general use		
TS-SPEA2	Alcalá et al. (2007d)	Tuning & Selection by SPEA2
TS-NSGA-II	Alcalá et al. (2007d)	Tuning & Selection by NSGA-II
TS-NSGA-II$_A$		Tuning & Selection by NSGA-II$_{\mathrm{angle}}$
TS-NSGA-II$_U$		Tuning & Selection by NSGA-II$_{\mathrm{utility}}$
Extended MOEAs for specific application		
TS-SPEA2$_{\mathrm{Acc}}$	Alcalá et al. (2007d)	Accuracy-Oriented SPEA2
TS-SPEA2$_{\mathrm{Acc}^2}$		Extension of SPEA2$_{\mathrm{Acc}}$

medium voltage lines in a town (1,059 cases; 4 continuous variables).

- The *Abalone* dataset (Waugh 1995) that concerns the task of trying to predict the number of rings in the shells of abalone (which is related to their age) based on a series of biometric measures of these animals (4,177 cases; 7 continuous variables; 1 nominal variable).

Methods considered for the experiments are briefly described in Table 2. In both problems, WM method is considered to obtain the initial set of fuzzy rules. To do so, we will consider symmetrical fuzzy partitions of triangular-shaped MFs. Once the initial rule set is generated, the proposed post-processing algorithms will be applied. T and S methods perform the tuning of parameters and rule selection respectively. TS indicates tuning together with rule selection in the same algorithm. All of them consider the accuracy of the model as the sole objective. MOEAs studied in this work (TS-SPEA2, TS-NSGA-II, TS-NSGA-II$_A$, TS-NSGA-II$_U$, TS-SPEA2$_{\mathrm{Acc}}$ and TS-SPEA2$_{\mathrm{Acc}^2}$) perform rule selection from a given fuzzy rule set together with the parametric tuning of the MFs considering two objectives, system error and number of rules.

In the next subsections, the named problems are introduced and solved to analyze the behavior of the proposed methods. To do that, the experimental set-up is first described. Finally, at the end of this section an internal study on alternative possibilities to select solutions from final Paretos and on the initialization influence is also performed (considering the initialization presented in Sect. 4.1 or a completely random initialization).

5.1 Experimental set-up

To develop the different experiments, we consider a *5-folder cross-validation model*, i.e., 5 random partitions of data each with 20%, and the combination of 4 of them (80%) as training and the remaining one as test. For each one of the five data partitions, the post-processing methods have been run 6 times (6 different seeds), showing for each problem the averaged results of a total of 30 runs.

In the case of methods with multi-objective nature (TS-SPEA2, TS-NSGA-II, TS-NSGA-II $_A$, TS-NSGA-II$_U$, TS-SPEA2$_{\mathrm{Acc}}$ and TS-SPEA2$_{\mathrm{Acc}^2}$), the averaged values are calculated considering the most accurate solution from each Pareto obtained. In this way, the multi-objective algorithms can be compared with several single objective based methods. This way to work differs with the previous works in the specialized literature (see Sect. 3) in which one or several Pareto fronts are presented and an expert should after select one solution. Our main aim following this approach is to compare the same technique when only the accuracy objective is considered (algorithm WM + TS) with the most accurate solution found by the proposed multi-objective algorithms in order to see if the Pareto fronts obtained are not only wide but also optimal (almost similar solutions to that obtained by WM + TS should be included in the final Pareto).

The values of the input parameters considered by S, T and TS (single objective oriented algorithms) are:[1] population size of 61, 100,000 evaluations, 0.6 as crossover probability and 0.2 as mutation probability per chromosome. In the case of MOEAs, the most important parameter is the population size. In the case of SPEA2 based algorithms, a good proportion between standard population and external population is 3/1 or 4/1. Different population sizes were probed showing not very different results but presenting the best performance when the external population took values between 50 and 100 individuals. It is, when the population used for parent selection has similar sizes than

[1] With these values we have tried to ease the comparisons selecting standard common parameters that work well in most cases instead of searching for very specific values to each method. Moreover, we have set a large number of evaluations in order to allow the compared algorithms to achieve an appropriate convergence.

those considered by single objective oriented algorithms in these kinds of problems. In this way, we have considered an external population size of 61 (the same size used by the named algorithms with single objective) and a proportion of 1/3 rounded to 200 as standard population size. In the case of NSGA-II based algorithms, since the archive and the population have the same size and they have been usually used with values of 200 and 100 as population size in general problems for continuous optimization, a good value could be 200 as population size (the same that SPEA2 based approaches). However, although 100, 200 and 400 presented a similar and reasonable performance for these algorithms, the best results were obtained by taking similar sizes than those considered by S, T and TS (with single objective) in these kinds of problems (i.e., when the size of the population used for parent selection takes these values). Therefore, we recommend the use of this simple rule of thumb to fix the population size in these kinds of prob lems. Finally, the values of the input parameters considered by the MOEAs are shown in the next: population size of 200 (61 in the case of NSGA-II based algorithms), external population size of 61 (in the case of SPEA2 based algorithms), 100000 evaluations and $P_m = 0.2$ as mutation probability per chromosome.

5.2 Estimating the maintenance costs of medium voltage lines

Estimating the maintenance costs of the medium voltage electrical network in a town (Cordón et al. 1999) is a complex but interesting problem. Since a direct measure is very difficult to obtain, it is useful to consider models. These estimations allow electrical companies to justify their expenses. Moreover, the model must be able to explain how a specific value is computed for a certain town. Our objective will be to relate the *maintenance costs of the medium voltage lines* with the following four variables: *sum of the lengths of all streets in the town*, *total area of the town*, *area that is occupied by buildings*, and *energy supply to the town*. We will deal with estimations of minimum maintenance costs based on a model of the optimal electrical network for a town in a sample of 1,059 towns. As said, five data partitions[2] considering an 80% (847) in training and a 20% (212) in test are considered for the experiments. The initial linguistic partitions are comprised by *five linguistic terms* with equally distributed triangular shaped MFs.

The results obtained by the analyzed methods are shown in Table 3, where *#R* stands for the number of rules, MSE_{tra} and MSE_{tst} respectively for the averaged error obtained over the training and test data, σ for the standard deviation and t for the results of applying a *student's t-test* (with 95 percent confidence) in order to ascertain whether differences in the performance of the best results are significant when compared with that of the other algorithms in the table. The interpretation of this column is:

* represents the best averaged result.
+ means that the best result has better performance than that of the corresponding row.

Analysing the results showed in Table 3 we can highlight the following facts:

- Methods based on SPEA2 show a reduction of MSE_{tra} and MSE_{tst} with respect to the models obtained by only considering the accuracy objective (WM + TS). Moreover, a considerable number of rules have been removed from the initial FRBS, obtaining simpler models with a better performance.
- NSGA-II based algorithms statistically obtain the same accuracy than the models obtained with TS-SPEA2 considering the most accurate result of each obtained Pareto. However, all of them present a higher number of rules (from one to three) and worse average values than TS-SPEA2. Moreover, a difference with TS-SPEA2 is that comparing each of the NSGA-II based approaches with WM + TS (single objective-based approach) the Student's t test would show that they are statistically equal in their generalization ability (MSE_{tst}). Therefore, we could consider that these algorithms get good solutions, since in any case, they are quite similar to the application of the standard SPEA2 (specially in the case of TS-NSGA-II$_A$ and TS-NSGA-II$_U$).
- The best results were obtained by TS-SPEA2$_{Acc^2}$ and TS-SPEA2$_{Acc}$, showing that the use of expert knowledge in the design process can help to obtain more

Table 3 Results obtained by the studied methods in the electrical distribution problem

Method	#R	MSE_{tra}	σ_{tra}	t	MSE_{tst}	σ_{tst}	t
WM	65.0	57,605	2,841	+	57,934	4,733	+
WM+T	65.0	17,020	1,893	+	21,027	4,225	+
WM+S	40.9	41,158	1,167	+	42,988	4,441	+
WM+TS	41.3	13,387	1,153	+	17,784	3,344	+
TS-SPEA2	**28.9**	11,630	1,283	+	15,387	3,108	=‡
TS-NSGA-II	31.4	11,826	1,354	+	16,047	4,070	+
TS-NSGA-II$_A$	29.7	11,798	1,615	+	16,156	4,091	+
TS-NSGA-II$_U$	30.7	11,954	1,768	+	15,879	4,866	+
TS-SPEA2$_{Acc}$	32.3	10,714	1,392	=	14,252	3,181	=
TS-SPEA2$_{Acc^2}$	29.8	**10,325**	**1,121**	*	**13,935**	**2,759**	*

Bold values represent the best results in each column

‡+ with 94% confidence

[2] These data sets are available at: http://decsai.ugr.es/~casillas/fmlib .

optimal Pareto fronts. Moreover, TS-SPEA2$_{Acc^2}$ is able to obtain the best average values with even less rules than the TS-SPEA2$_{Acc}$ algorithm.

All MOEAs considered obtain significantly simpler models that those obtained by only considering the accuracy based objective and almost the same results (presenting minor average values in all the cases and statistical differences in the case of the extended MOEAs). This is a positive fact since an appropriate use of MOEAs can improve the desired trade-off with respect to the classic accuracy-based approaches, and specific adaptations can help to improve the performance of standard MOEAs.

These results (more simple and accurate models by applying a multi-objective approach) are due to the large search space that involves these kinds of problems. There are some initial rules that should be removed since they do not cooperate in a good way with the remaining ones. Even in the case of only considering an accuracy-based objective, the large search space that supposes the tuning of parameters makes very difficult to remove these kinds of rules since bad rules are tuned together with the remaining ones searching for their best cooperation. The use of a multi-objective approach favors a better selection of the ideal number of rules, preserving some rule configurations until the rule parameters are evolved to dominate solutions including bad rules, which can finally lead to solutions with more freedom-degrees to tune the corresponding parameters involving a better cooperation among the different rules.

In Fig. 3, we can see the Pareto evolution in a representative run for each multi-objective algorithm and also the evolution of the best solution in the population in a representative run of WM+TS. Each type of symbol in the figure represents the Pareto solutions at different stages of the evolution (caption 'Evaluations' shows the number of evaluations in which each Pareto was taken in a simple run and the symbol associated). We can observe as the Pareto moves along without having a wide extension but

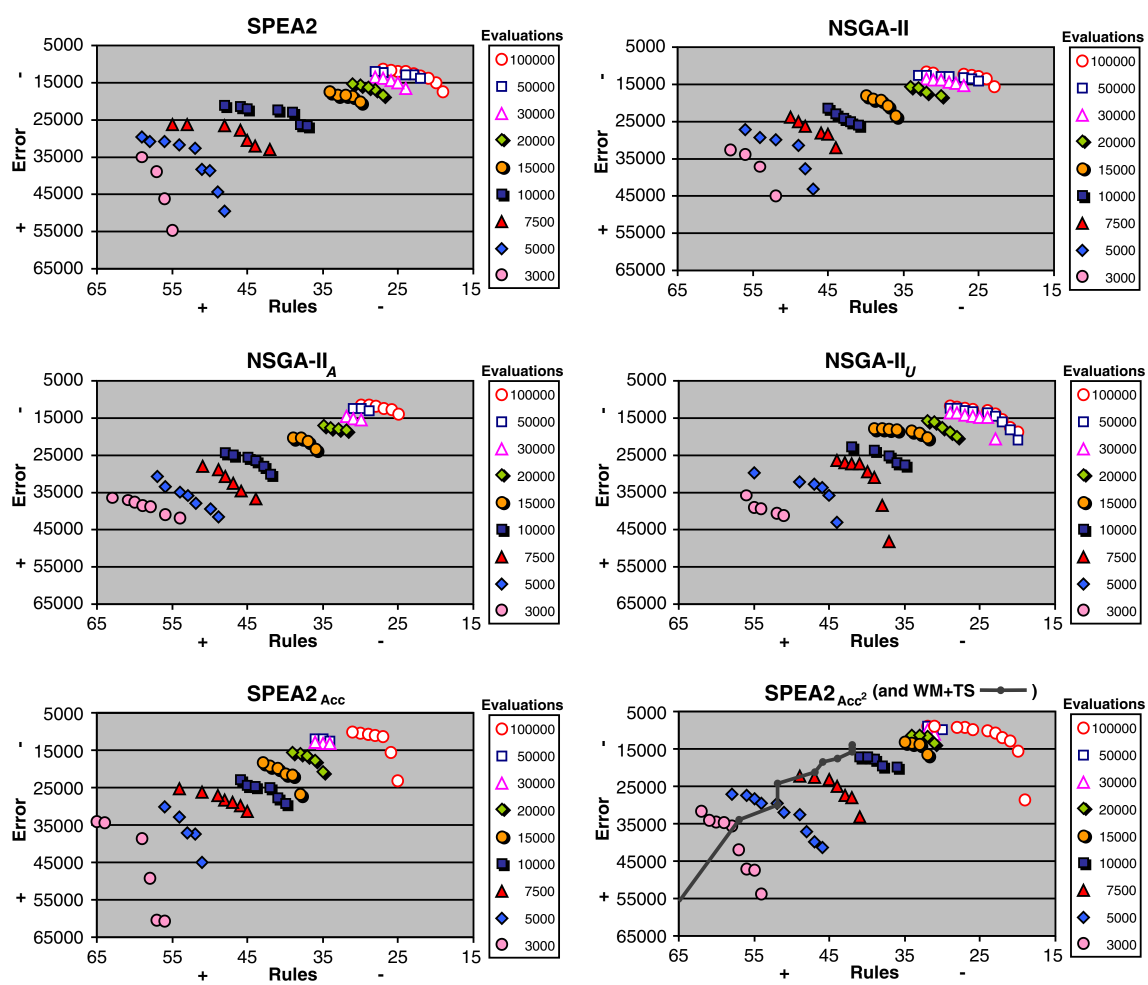


Fig. 3 Example of the Pareto front evolution along one representative run of TS-SPEA2, TS-NSGA-II, TS-NSGA-II$_A$, TS-NSGA-II$_U$, TS-SPEA2$_{Acc}$ and TS-SPEA2$_{Acc^2}$ in the Electrical distribution problem. Evolution of the best solution in a representative run with WM+TS is also included together with TS-SPEA2$_{Acc^2}$

dominating the solution obtained by WM+TS at the end. Although these figures only represent a run for each algorithm, the obtained Pareto fronts in the different runs (5 fold, six seeds, 30 runs, after 100,000 evaluations) are in general very similar to those showed in Fig. 3 for each algorithm. In this way, the most accurate solution of each Pareto can be considered as the position in which we can find a sort set of close solutions representing different trade-offs in the Pareto zone with still accurate solutions.

Another important fact is that in the final Pareto fronts of TS-SPEA2$_{Acc^2}$ and TS-SPEA2$_{Acc}$ there are one or two solutions (those with the minor number of rules) showing a bad performance with respect to the remaining ones in their respective Paretos. This situation is typical in practically all the obtained Pareto fronts and in all the approaches considered. These solutions represent new individuals that appeared at the end of the evolution practically without time to evolve their associated MFs in a zone in which to extract a rule without severely affecting the accuracy is more difficult.

Figure 4 shows the convergence of the best solution of the population from WM+TS and the most accurate solution in the Pareto from TS-SPEA2$_{Acc^2}$ in a representative run for the electrical distribution problem. An interesting fact is that the WM+TS (single objective) algorithm is faster at the beginning (in training and number of rules) while TS-SPEA2$_{Acc^2}$ takes a bit more time for exploration in order to take further advantage. Another interesting fact is to see how to perform restarting at the mid of the TS-SPEA2$_{Acc^2}$ process positively affects the values in training and number of rules.

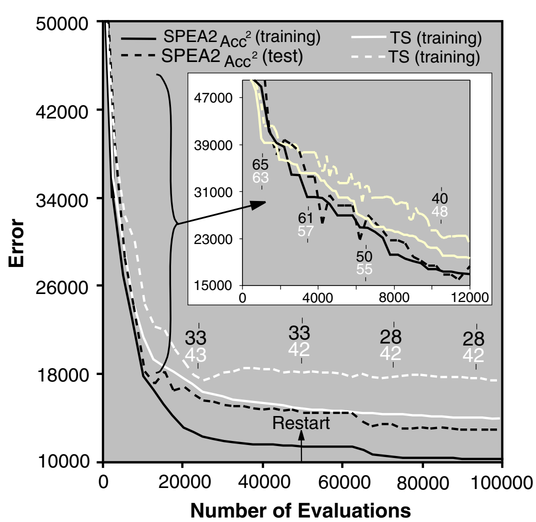


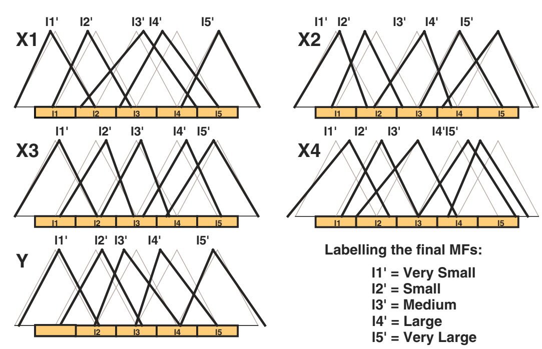
Fig. 4 Convergence in a representative run considering the best solution of the population of WM+TS and the most accurate solution in the Pareto of TS-SPEA2$_{Acc^2}$. Numbers in black and white represent the number of rules in the solution in different moments of the evolution

Figures 5 and 6, respectively show the most accurate models obtained with WM+TS and TS-SPEA2$_{Acc^2}$ in the electrical distribution problem. To ease graphic representation, in these figures, the MFs are labeled from $l1$ to lm^i. Nevertheless, such MFs are initially associated to a linguistic meaning determined by an expert. In this way, if the $l1$ label of the $X1$ variable represents 'Very Small', $l1'$ could be also interpreted as 'Very Small' as classically has been considered when tuning is applied or based on the expert opinion, maintaining the original meaning of such label or renaming it if possible. This is the case of Figs. 5 and 6, where in principle practically all the new labels could maintain their initial meanings.

5.3 Predicting the Abalone age

The Abalone dataset (Waugh 1995) is concerned with predicting the age of an Abalone specimen (a type of shellfish) based on physical measurements. Why is it interesting to predict age? For ecologic and commercial fish farming purposes, the age composition of abalone populations are relevant. Here the number of rings is proxy for age. The age of abalone is determined by cutting the shell through the cone, staining it, and counting the number of rings through a microscope.

However, this is a boring, time-consuming and expensive task. Other measurements, which are easier to obtain, are therefore used to predict the age. Recorded measurements on



#R: 43 *MSE-tra: 11383* *MSE-tst: 13416*

X1	X2	X3	X4	Y
l1'	l1'	l1'	l1'	l1'
l1'	l1'	l1'	l2'	l2'
l1'	l2'	l1'	l1'	l1'
l1'	l2'	l2'	l2'	l2'
l2'	l1'	l1'	l1'	l1'
l2'	l1'	l1'	l2'	l2'
l2'	l1'	l2'	l1'	l2'
l2'	l1'	l2'	l2'	l2'
l2'	l2'	l1'	l1'	l1'
l2'	l2'	l1'	l2'	l2'
l2'	l2'	l2'	l1'	l2'
l2'	l2'	l2'	l2'	l2'
l2'	l3'	l3'	l2'	l3'
l3'	l2'	l1'	l1'	l1'

X1	X2	X3	X4	Y
l3'	l2'	l1'	l2'	l2'
l3'	l2'	l1'	l3'	l2'
l3'	l2'	l2'	l3'	l3'
l3'	l3'	l2'	l1'	l2'
l3'	l3'	l2'	l2'	l2'
l3'	l3'	l2'	l3'	l3'
l3'	l3'	l3'	l2'	l3'
l3'	l4'	l3'	l2'	l3'
l3'	l4'	l4'	l3'	l4'
l4'	l2'	l2'	l1'	l2'
l4'	l2'	l2'	l2'	l2'
l4'	l3'	l2'	l1'	l2'
l4'	l3'	l2'	l3'	l3'
l4'	l3'	l2'	l4'	l3'

X1	X2	X3	X4	Y
l4'	l3'	l3'	l2'	l3'
l4'	l3'	l3'	l3'	l4'
l4'	l4'	l3'	l1'	l3'
l4'	l4'	l3'	l4'	l4'
l4'	l4'	l4'	l2'	l4'
l4'	l4'	l4'	l4'	l5'
l4'	l5'	l4'	l2'	l3'
l4'	l5'	l4'	l3'	l4'
l4'	l5'	l5'	l2'	l5'
l4'	l5'	l5'	l3'	l5'
l5'	l2'	l2'	l5'	l4'
l5'	l2'	l3'	l2'	l3'
l5'	l2'	l3'	l5'	l4'
l5'	l4'	l3'	l4'	l4'
l5'	l4'	l3'	l5'	l5'

Fig. 5 DB with/without tuning (*black/gray*) and RB of the best model (in training) obtained by WM+TS in the electrical distribution problem

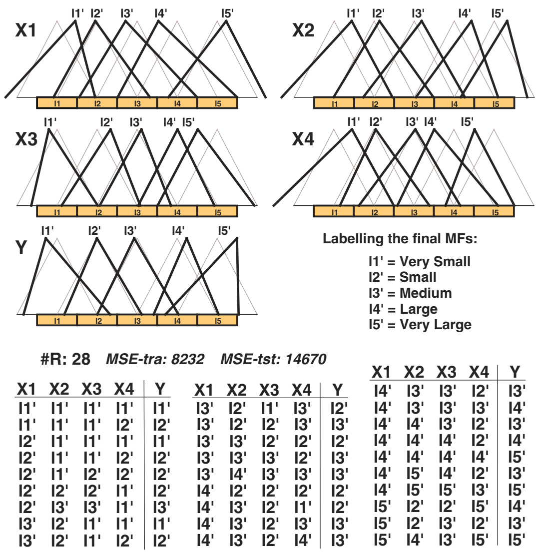


Fig. 6 DB with/without tuning (*black/gray*) and RB of the best model (in training) obtained by TS-SPEA2$_{\mathrm{Acc}^2}$ in the electrical distribution problem

4,177 Abalone,[3] of interest in determining relationships useful to predicting the age of future abalone from easily made physical measurements, were obtained from the Marine Resources Division at the Department of Primary Industry and Fisheries, Tasmania, and can be used for this task. The goal is to predict the *number of rings* based on the following eight variables (seven continuous and one nominal): *sex* (nominal), *length*, *diameter*, *height*, *whole weight*, *shucked weight*, *viscera weight*, and *shell weight*. As explained, five data partitions considering an 80% (3,342) in training and a 20% (835) in test are considered for the experiments. In this case, the initial linguistic partitions are comprised by *three linguistic terms* with equally distributed triangular shaped MFs. The accuracy of the models obtained is quite similar by considering three or five linguistic terms in this problem and therefore a number of three labels per variable is preferable since the final models are comprised of a smaller number of rules.

The results obtained in this problem by the analyzed methods are shown in Table 4 (these kinds of table was described in the previous subsection). On the Abalone data set (with very high level of noise), all of the MOEAs and even WM + TS achieved almost identical performance, however they present different numbers of rules. This problem is quite different to that in the previous section. As can be seen, WM+S obtains the models with the smaller number of rules, which indicates that there are a lot of rules that can be removed from the initial system. So, in this problem, the real challenge would be to remove these rules

[3] Available from the UCI Machine Learning Repository (http://www.ics.uci.edu/~mlearn/MLRepository.html).

Table 4 Results obtained by the studied methods in the Abalone dataset

Method	#R	MSE$_{tra}$	σ_{tra}	t	MSE$_{tst}$	σ_{tst}	t
WM	68.2	8.407	0.443	+	8.422	0.545	+
WM+T	68.2	2.688	**0.063**	+	2.770	0.242	+
WM+S	**18.0**	4.825	1.078	+	4.795	1.165	+
WM+TS	28.4	2.473	0.097	+	2.582	0.290	=
TS-SPEA2	20.0	2.383	0.078	=	2.518	0.246	=
TS-NSGA-II	22.4	2.398	0.084	=	2.526	0.242	=
TS-NSGA-II$_A$	22.1	2.404	0.098	=	2.535	0.265	=
TS-NSGA-II$_U$	21.8	2.407	0.082	=	2.520	0.237	=
TS-SPEA2$_{\mathrm{Acc}}$	22.2	**2.368**	0.085	*	**2.511**	0.263	*
TS-SPEA2$_{\mathrm{Acc}^2}$	18.6	2.372	0.075	=	2.517	**0.230**	=

Bold values represent the best results in each column

in an appropriate manner instead of trying very important accuracy improvements. Analyzing the results presented in Table 4 we can stress the following facts:

- In this case, although TS-SPEA2$_{\mathrm{Acc}}$ method presents the best average results in MSE$_{tra}$ and MSE$_{tst}$ with respect to the remaining models, TS-SPEA2$_{\mathrm{Acc}^2}$ could be considered as the best approach since practically the same values were obtained in training and test, and the best value in number of rules has been obtained.
- In this case, NSGA-II based algorithms are statistically equal than those models obtained by application of the standard SPEA2, but again, all of them present a higher number of rules (about 2). However, accuracy differences are practically not appreciated showing results quite similar to TS-SPEA2. This time, TS-NSGA-II$_A$ and TS-NSGA-II$_U$ obtain more or less the same number of rules than TS-NSGA-II, although their average numbers of rules are still better.
- All MOEAs considered obtain significantly simpler models that those obtained by only considering the accuracy based objective and almost the same results, improving again the desired trade-off with respect to the classic accuracy-based approaches.

In Fig. 7, we can see the Pareto evolution in a representative run for each multi-objective algorithm (these kinds of figures were described in the previous subsection). Once more, the different Pareto fronts move along without having a wide extension. TS-NSGA-II$_U$, TS-SPEA2$_{\mathrm{Acc}}$ and TS-SPEA2$_{\mathrm{Acc}^2}$ show the wider Pareto fronts in their corresponding figures. However, the front obtained by TS-SPEA2$_{\mathrm{Acc}^2}$ is again located more to the top right zone (the zone with less rules and more accurate models).

Figure 8 shows the most accurate model obtained with TS-SPEA2$_{\mathrm{Acc}^2}$ in the electrical distribution problem (these

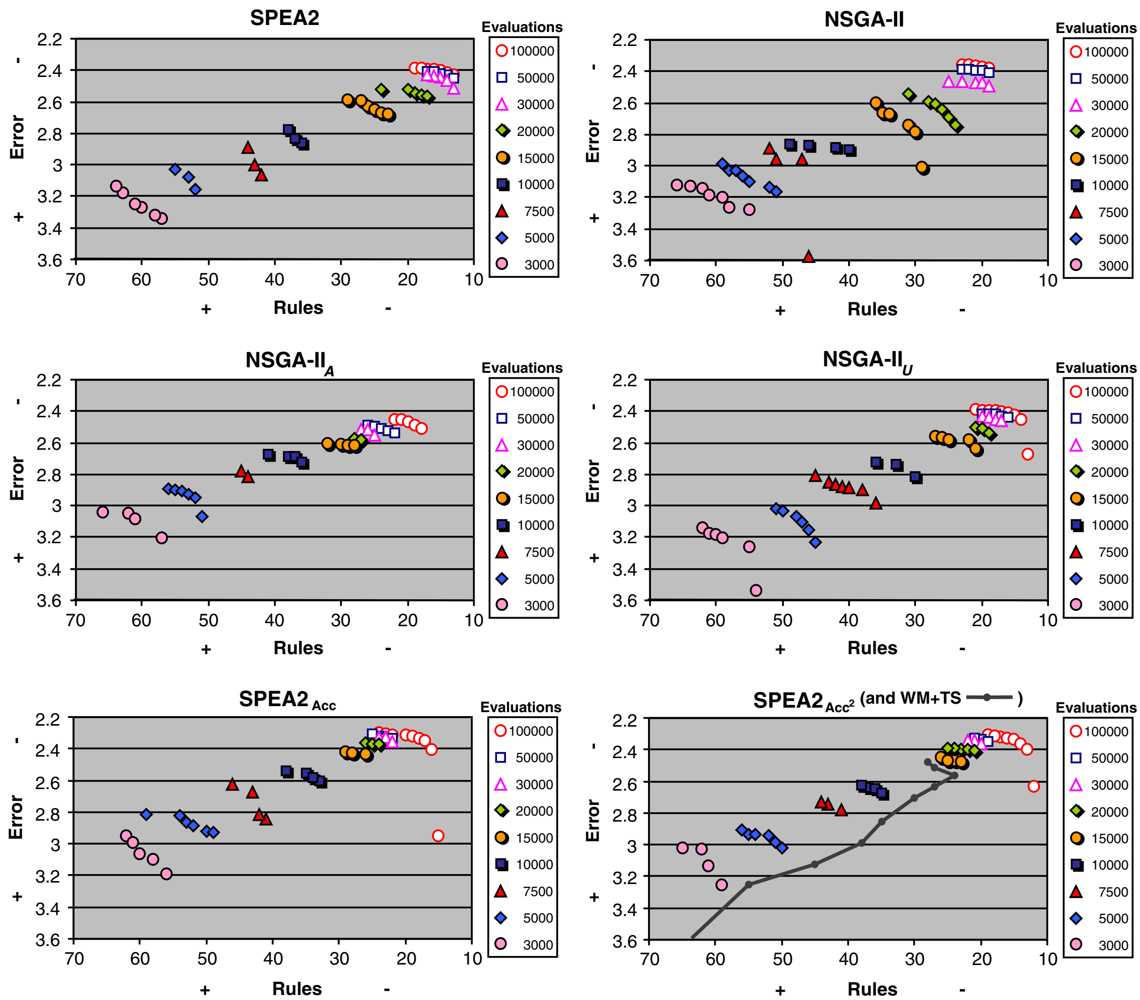


Fig. 7 Example of the Pareto front evolution along one representative run of TS-SPEA2, TS-NSGA-II, TS-NSGA-II$_A$, TS-NSGA-II$_U$, TS-SPEA2$_{Acc}$ and TS-SPEA2$_{Acc^2}$ in the Abalone dataset. Evolution of the best solution in a representative run with WM+TS is also included together with TS-SPEA2$_{Acc^2}$

kinds of figures were also described in the previous subsection). Although practically all the MFs could maintain their original meanings (from a subjective point of view), there are two cases that probably should be renamed by experts if possible, l2' and l3' in X1 and X6, respectively. From a subjective point of view we show a way to rename them. The most accurate model obtained from WM+TS has been not included for this problem in order to avoid an excessive length of the paper since, as in the previous problem, its MFs are similar to those obtained by TS-SPEA2$_{Acc^2}$ (even with any MFs that should be renamed).

5.4 Analysis on the solution selection and importance of the initialization

As said, an internal study on alternative possibilities to select solutions from final Pareto fronts and on the initialization influence is also performed by focusing on the electrical problem (the one with more possibilities to reduce not only the rule number but also the system error). Although we propose as final solution the most accurate one since our main objective is to reduce the number of rules but maintaining or improving the accuracy of the obtained models, there is another motivation that reinforced our decision. If this solution is maintained as a part of the final Pareto is because no other rule configuration is able to obtain a better parameter tuning (the main reason of the improved accuracy of these kinds of models). So we can be sure that this solution had the time to be evolved more or less in the proper way. However, the more we look for simpler solutions in the final Pareto the less we can be sure that these solutions had the time to be tuned (if a simpler solution appears at the end of the evolutionary process probably this solution had not the time to be

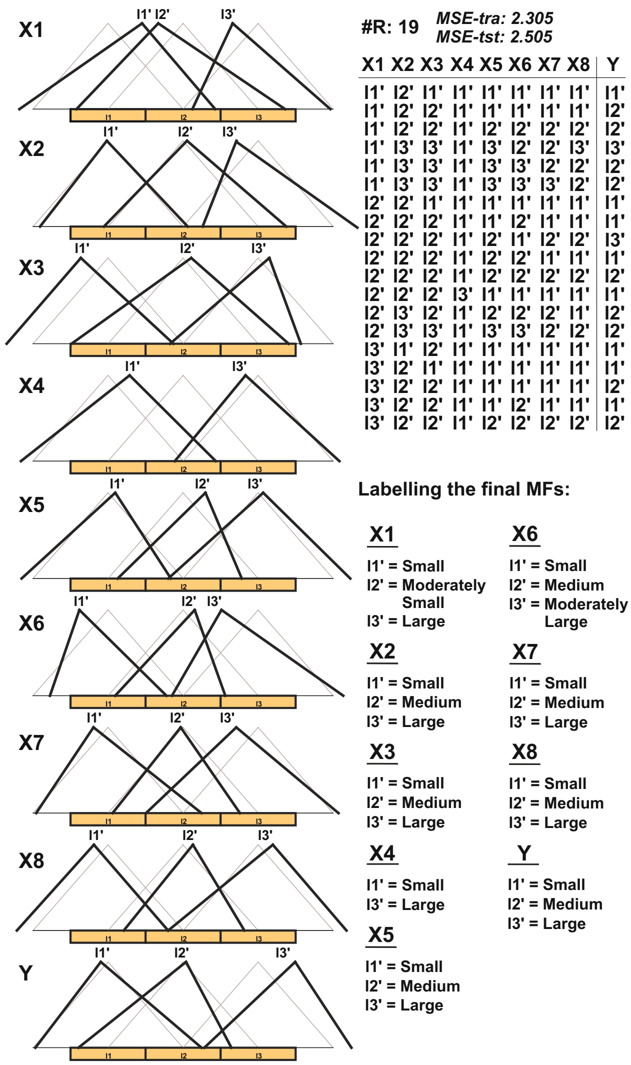


Fig. 8 DB with/without tuning (*black/gray*) and RB of a model obtained by TS-SPEA2$_{\mathrm{Acc}^2}$ in the Abalon dataset

properly tuned). In any case, the output of MOEAs really is a set of solutions from which an expert could choose the most convenient one.

In Table 5, we consider two different possibilities applied on the results obtained in Sect. 5.2 with TS-NSGA-II$_A$, TS-NSGA-II$_U$ and TS-SPEA2$_{\mathrm{Acc}^2}$. In this case, we choose the solution with the best angle or utility measure in the Pareto fronts obtained from TS-NSGA-II$_A$ and TS-NSGA-II$_U$, respectively, and the i-th most accurate solution in the Pareto fronts obtained from TS-SPEA2$_{\mathrm{Acc}^2}$. As can be seen, the results obtained by choosing the proper i-th solution from TS-SPEA2$_{\mathrm{Acc}^2}$ outperform the results obtained by knee based approaches. Since a knee is an ideal solution in the ideal Pareto but a promising one in a population of non dominated solutions, the knee based measure can be good to favor the evolution of promising

Table 5 Results obtained by choosing the knee or the i-th most accurate solution in the electrical distribution problem

Method	#R	MSE$_{tra}$	σ_{tra}	t	MSE$_{tst}$	σ_{tst}	t
TS-NSGA-II$_A$-k	25.5	13,242	2,383	+	17,541	4,184	+
TS-NSGA-II$_U$-k	24.2	15,797	3,945	+	20,528	6,802	+
TS-SPEA2$_{\mathrm{Acc}^2}$ −1	29.8	**10,325**	**1,121**	*	**13,935**	**2,759**	*
TS-SPEA2$_{\mathrm{Acc}^2}$ −2	28.3	10,496	1,126	=	14,268	2,925	=
TS-SPEA2$_{\mathrm{Acc}^2}$ −3	27.0	10,835	1,191	=	14,460	2,782	=
TS-SPEA2$_{\mathrm{Acc}^2}$ −4	25.9	11,217	1,307	+	14,806	3,069	=
TS-SPEA2$_{\mathrm{Acc}^2}$ −5	24.9	12,194	2,078	+	15,417	3,328	=

Bold values represent the best results in each column

solutions but it seems not so good to choose the final solution from evolved Pareto fronts that could present false knees (for example one solution with too few rules but without time to be properly tuned can be identified as a knee). In fact, standard deviations in the table show the diversity of solutions proposed considering knee measures (different knees appears in different runs, and for example a solution with 20 rules 27,426 in training and 38432 in test is proposed by TS-NSGA-II$_U - k$ in one of the 30 runs). In any case, independently of the mechanism considered to propose a final solution the most important thing is if the obtained front can be nearer of the optimal one, and the situation of the most accurate solution seems to be very indicative in this sense.

A study has been also performed on the importance of the initialization component for the rule selection part in the chromosome (considering a completely random initialization instead of the one presented in Section 4.1). Table 6 presents the results (again considering the most accurate solution in the final fronts) obtained by all the MOEAs considered for comparison starting with the initial rules selected at random (the best result in Table 3 is also included to show differences). By considering random initialization the results obtained present too low numbers of rules with much worse results especially in the test. This shows that to add rules that were not selected at the beginning is not easy since the MF parameters are quickly adapted to those rules that are just selected giving way to sub-optimal Pareto fronts. In any case, there are two important facts in these results:

- TS-SPEA2$_{\mathrm{Acc}}$ and TS-SPEA2$_{\mathrm{Acc}^2}$ methods were not very affected by the random initialization, presenting solutions that were also interesting from the trade-off point of view (very low number of rules and a good accuracy).
- Fixing the number of rules in the initial population can be a way to regulate the desired trade-off since this biases the number of rules in the final solutions.

Table 6 Results obtained by random initialization of the C_S part (rule part) in the electrical distribution problem

Method	#R	MSE_{tra}	σ_{tra}	t	MSE_{tst}	σ_{tst}	t
TS-SPEA2	21.9	18,768	2,256	+	23,951	**5,198**	+
TS-NSGA-II	27.7	17,688	2,333	+	23,762	7,681	+
TS-NSGA-II$_A$	**19.5**	23,981	3,709	+	29,442	7,058	+
TS-NSGA-II$_U$	24.4	18,728	2,071	+	24,148	5,397	+
TS-SPEA2$_{Acc}$	23.2	**14,175**	**1,752**	*	**18,289**	5,571	*
TS-SPEA2$_{Acc^2}$	20.2	16,539	2,729	+	21,977	5,625	+
Best in Table 3:	29.8	10,325	1,121		13,935	2,759	

Bold values represent the best results in each column

6 Concluding remarks

In this work, we have analyzed the application of different MOEAs to obtain simpler but still accurate linguistic fuzzy models by performing rule selection and a classic tuning of the MF parameters. In order to show the main differences with the previous works, a brief analysis of the state of the art on the use of MOEAs to get FRBSs with good accuracy-interpretability trade-off has been performed at first. From this study we can stress the following points:

- Most of the works only consider quantitative measures of the system complexity to determine the FRBS interpretability since the use of qualitative measures is still an open topic that needs of further and intense research efforts.
- None of the works (but the one in Alcalá et al. (2007d)) considered a learning or tuning of the MFs, only performing rule learning or selection.
- Algorithms considered were slight modifications of MOEAs proposed for general use (MOGA, NSGA-II, etc.) or specifically developed for this concrete and difficult problem. It is due to the special nature of this problem, in which to improve the accuracy objective is more difficult than simplifying the fuzzy models, by which the Pareto front finally obtained still becomes sub-optimal with respect to the accuracy objective. Therefore, MOEAs considering specific information about the problem are usually needed.

Since combining rule selection and tuning of the system parameters represents a more complex search space and therefore needs of different considerations with respect to the works in the existing literature, some considerations based on the experience are needed in the MOEA design process in order to get good solutions. From the results obtained, we can conclude that:

- The results obtained have shown that an appropriate use of MOEAs can represent a way to obtain even more accurate and simpler linguistic models than those obtained by only considering performance measures.
- Population initialization is an important component that can help to regulate the desired trade-off since this biases the number of rules in the final solutions. In this way, the results obtained by selecting all the rules in the initial population are able to find solutions in the most accurate Pareto zone.
- The best results were obtained by TS-SPEA2$_{Acc^2}$ on two different scenarios. These results show that the use of experience based knowledge in the MOEAs design process can significantly improve the search ability of these algorithms.

Finally, we would like to point out that the analysis presented in this work could help to extend this approach in order to consider other kinds of techniques or new interpretability measures for further works, e.g., another tuning types, learning, combination with the use of quality measures, etc.

Acknowledgment Supported in part by the Spanish Ministry of Education and Science under grant no. TIN2005-08386-C05-01, and the Andalusian government under grant no. P05-TIC-00531.

References

Alcalá R, Alcalá-Fdez J, Casillas J, Cordón O, Herrera F (2006) Hybrid learning models to get the interpretability-accuracy trade-off in fuzzy modeling. Soft Comput 10(9):717–734

Alcalá R, Alcalá-Fdez J, Herrera F, Otero J (2007a) Genetic learning of accurate and compact fuzzy rule based systems based on the 2-tuples linguistic representation. Int J Approx Reason 44(1):45–64

Alcalá R, Alcalá-Fdez J, Gacto MJ, Herrera F (2007b) Rule base reduction and genetic tuning of fuzzy systems based on the linguistic 3-tuples representation. Soft Computing 11(5):401–419

Alcalá R, Alcalá-Fdez J, Herrera F (2007c) A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection. IEEE Trans Fuzzy Syst 15(4):616–635

Alcalá R, Gacto MJ, Herrera F, Alcalá-Fdez J (2007d) A multi-objective genetic algorithm for tuning and rule selection to obtain accurate and compact linguistic fuzzy rule-based systems. Int J Uncertain Fuzziness Knowl Based Syst 15(5):539–557

Branke J, Deb K, Dierolf H, Osswald M (2004) Finding knees in multi-objective optimization. In: Proc Parallel Problem Solving from Nature Conf.—PPSN VIII. LNCS, vol 3242. Birmingham, UK, pp 722–731

Casillas J, Cordón O, Herrera F, Magdalena L (eds) (2003a) Interpretability issues in fuzzy modeling. Studies in Fuzziness and Soft Computing. Springer, Heidelberg, pp 128

Casillas J, Cordón O, Herrera F, Magdalena L (eds) (2003b) Accuracy improvements in linguistic fuzzy modelling. Studies in Fuzziness and Soft Computing. Springer, Heidelberg, pp 129

Casillas J, Cordón O, del Jesus MJ, Herrera F (2005) Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction. IEEE Trans Fuzzy Syst 13(1):13–29

Cococcioni M, Ducange P, Lazzerini B, Marcelloni F (2007) A Pareto-based multi-objective evolutionary approach to the identification of mamdani fuzzy systems. Soft Comput 11:1013–1031

Coello CA, Toscano G (2001) A Micro-Genetic Algorithm for multiobjective optimization. In: First international conference on evolutionary multi-criterion optimization. LNCS 1993, London, UK, pp 126–140

Coello CA, Van Veldhuizen DA, Lamont GB (2002) Evolutionary algorithms for solving multi-objective problems. Kluwer, Dordrecht

Cordón O, Herrera F, Sánchez L (1999) Solving electrical distribution problems u-sing hybrid evolutionary data analysis techniques. Appl Intell 10:5–24

Cordon O, Herrera F, del Jesus MJ, Villar P (2001) A multiobjective genetic algorithm for feature selection and granularity learning in fuzzy-rule based cla-ssi-fication systems. In: Proceedings of IX IFSA World Congress and XX NAFIPS Int. Conf., Vancouver, Canada, pp 1253–1258

Corne D, Knowles J, Oates M (2000) The Pareto Envelope-based Selection Algorithm for multiobjective optimization. In: Proc. Parallel Problem Solving from Nature Conf.—PPSN VI, LNCS 1917. Paris, France, pp 839–848

Corne D, Jerram N, Knowles J, Oates M (2001) PESA-II: Region based selection in evolutionary multiobjective optimization. In: Proceedings of genetic and evolutionary computation conf., San Francisco, pp 283–290

Deb K (2001) Multi-objective optimization using evolutionary algorithms. Wiley, New York

Deb K, Agrawal S, Pratab A, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evolut Comput 6(2):182–197

Erickson M, Mayer A, Horn J (2001) The Niched Pareto Genetic Algorithm 2 a-pplied to the design of groundwater remediation systems. In: Proceedings of first international conference on evolutionary multi-criterion optimization. LNCS, vol 1993. London, UK, pp 681–695

Eshelman LJ (1991) The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. Found Genetic Algorithms 1:265–283

Eshelman LJ, Schaffer JD (1993) Real-coded genetic algorithms and interval-schemata. Found Genetic Algorithms 2:187–202

Fonseca CM, Fleming PJ (1993) Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. In: Proceedings of 5th international conference on genetic algorithms, San Mateo, pp 416–423

Goldberg D (2000) A meditation on the computational intelligence and its future. Illigal Report #2000019, Department of General Engineering, University of Illinois at Urbana-Champaign, Foreword Proc. of the 2000 Int. Symposium on Computational Intelligence

Horn J, Nafpliotis N, Goldberg DE (1994) A niched Pareto genetic algorithm for multiobjective optimization. In: Proceedings of first IEEE conference on evolutionary computation, IEEE World Congress on Computational Intelligence, vol 1, Piscataway, pp 82–87

Ishibuchi H, Murata T (1996) Multi-objective genetic local search algorithm. In: Proceedings of third IEEE international conference on evolutionary computation, Japan, pp 119–124

Ishibuchi H, Nojima Y (2007) Analysis of interpretability-accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning. Int J Approx Reason 44(1):4–31

Ishibuchi H, Yamamoto T (2003) Interpretability issues in fuzzy genetics-based machine learning for linguistic modelling. In: Lawry J, Shanahan JG, Ralescu AL (ed) Modelling with words: learning, fusion, and reasoning within a formal liguistic representation framework. LNCS, vol 2873. Springer, Berlin, pp 209–228

Ishibuchi H, Yamamoto T (2004) Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining. Fuzzy Sets Syst 141(1):59–88

Ishibuchi H, Murata T, Turksen IB (1997) Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems. Fuzzy Sets Syst 89(2):135–150

Ishibuchi H, Nakashima T, Murata T (2001) Three-objective genetics-based machine learning for linguistic rule extraction. Inf Sci 136:109–133

Jimenez F, Gomez-Skarmeta AF, Roubos H, Babuska R (2001) Accurate, transparent, and compact fuzzy models for function approximation and dynamic modeling through multi-objective evolutionary optimization. In: Proceedings of first international conference on evolutionary multi-criterion optimization. LNCS, vol 1993, Zurich, Switzerland, pp 653–667

Knowles JD, Corne DW (2000) Approximating the non dominated front using the Pareto archived evolution strategy. Evolut Comput 8(2):149–172

Narukawa K, Nojima Y, Ishibuchi H (2005) Modification of evolutionary multiobjective optimization algorithms for multi-objective design of fuzzy rule-based classification systems. In: Proceedings of 2005 IEEE international conference on fuzzy systems, Reno, USA, pp 809–814

Mamdani EH, Assilian S (1975) An experiment in linguistic synthesis with a fuzzy logic controller. Int J Man Mach Stud 7:1–13

Rosenberg RS (1967) Simulation of genetic populations with biochemical properties. MA Thesis, Univ. Michigan, Ann Harbor, Michigan

Schaffer JD (1985) Multiple objective optimization with vector evaluated genetic algorithms. In: Proceedings of first international conference on genetic algorithms, Pittsburgh, pp 93–100

Srinivas N, Deb K (1994) Multiobjective optimization using nondominated sorting in genetic algorithms. Evolut Comput 2:221–248

Takagi T, Sugeno M (1985) Fuzzy identification of systems and its application to modeling and control. IEEE Trans Syst Man Cybern 15(1):116–132

Wang LX, Mendel JM (1992) Generating fuzzy rules by learning from examples. IEEE Trans Syst Man Cybern 22(6):1414–1427

Wang HL, Kwong S, Jin YC, Wei W, Man KF (2005a) Multi-objective hierarchical genetic algorithm for interpretable fuzzy rule-based knowledge extraction. Fuzzy Sets Syst 149(1):149–186

Wang HL, Kwong S, Jin YC, Wei W, Man KF (2005b) Agentbased evolutionary approach for interpretable rule-based knowledge extraction. IEEE Trans Syst Man Cybern C Appl Rev 35(2):143–155

Waugh S (1995) Extending and Benchmarking Cascade-Correlation. PhD Thesis, Computer Science Department, University of Tasmania

Zitzler E, Thiele L (1999) Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. IEEE Trans Evolut Comput 3(4):257–271

Zitzler E, Laumanns M, Thiele L (2001) SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization. In: Proc evolutionary methods for design, optimization and control with app to industrial problems, Barcelona, Spain, pp 95–100

4. Integración de un Índice para Preservar la Interpretabilidad Semántica en la Selección y Ajuste Evolutivos Multi-Objetivo de los Sistemas Difusos Lingüísticos - *Integration of an Index to Preserve the Semantic Interpretability in the Multi-Objective Evolutionary Rule Selection and Tuning of Linguistic Fuzzy Systems*

Las publicaciones en revista asociadas a esta parte son:

- M.J. Gacto, R. Alcalá, F. Herrera, Integration of an Index to Preserve the Semantic Interpretability in the Multi-Objective Evolutionary Rule Selection and Tuning of Linguistic Fuzzy Systems. IEEE Transactions on Fuzzy Systems 18:3 (2010) 515-531, doi:10.1109/TFUZZ.2010.2041008.
    - Estado: Publicado.
    - Índice de Impacto (JCR 2009): 3,343.
    - Área de Conocimiento: Computer Science, Artificial Intelligence. Ranking 6 / 102.
    - Área de Conocimiento: Engineering, Electrical & Electronic. Ranking 10 / 245.

Integration of an Index to Preserve the Semantic Interpretability in the Multiobjective Evolutionary Rule Selection and Tuning of Linguistic Fuzzy Systems

María José Gacto, Rafael Alcalá, and Francisco Herrera

***Abstract*—In this paper, we propose an index that helps preserve the semantic interpretability of linguistic fuzzy models while a tuning of the membership functions (MFs) is performed. The proposed index is the aggregation of three metrics that preserve the original meanings of the MFs as much as possible while a tuning of their definition parameters is performed. Additionally, rule-selection mechanisms can be used to reduce the model complexity, which involves another important interpretability aspect. To this end, we propose a postprocessing multiobjective evolutionary algorithm that performs rule selection and tuning of fuzzy-rule-based systems with three objectives: accuracy, semantic interpretability maximization, and complexity minimization. We tested our approach on nine real-world regression datasets. In order to analyze the interaction between the fuzzy-rule-selection approach and the tuning approach, these are also individually proved in a multiobjective framework and compared with their respective single-objective counterparts. We compared the different approaches by applying nonparametric statistical tests for pairwise and multiple comparisons, taking into consideration three representative points from the obtained Pareto fronts in the case of the multiobjective-based approaches. Results confirm the effectiveness of our approach, and a wide range of solutions is obtained, which are not only more interpretable but are also more accurate.**

***Index Terms*—Fuzzy-rule-based systems (FRBSs), multiobjective evolutionary algorithms (MOEAs), rule selection, semantic interpretability index, tuning.**

I. Introduction

As DISCUSSED by Zadeh [1], computing with words (CW) is a methodology in which the objects of computation are words and propositions that are drawn from a natural language, e.g., small, large, far, etc. CW is inspired by the remarkable human capability to perform a wide variety of physical and mental tasks with no measurements or computations. CW-based techniques are employed to translate propositions that are expressed in a natural language into the generalized constraint language. The development of the methodology of CW is the development of a methodology in which words play the role of labels of perceptions. Linguistic variables and linguistic fuzzy rules are important elements in the conceptual structure of computational theory of perceptions (see [1, Fig. 4]). Further, as Zadeh stated [1], a fuzzy rule can be considered to be a Cartesian granule, and a fuzzy graph or a rule base (RB) may be viewed as a disjunction of Cartesian granules, and in essence, a fuzzy graph serves as an approximation to a function or a relation. This way, linguistic fuzzy modeling allows the modeling of systems to be dealt with by building a linguistic model that is interpretable by human beings. This task is usually developed by means of linguistic fuzzy-rule-based systems (FRBSs), which are also called Mamdani FRBSs [2], [3] and use fuzzy rules composed of linguistic variables [4]–[6] that take values in a term set with a real-world meaning, i.e., a variable whose values are words drawn from a natural language that represents the basis for the concept of linguistic *if–then* rules.

Many automatic techniques have been proposed to extract a proper set of linguistic fuzzy rules from numerical data. Most of them usually try to improve the performance that is associated with the prediction error without paying special attention to system's interpretability and without losing the linguistic meanings associated with the model. Finding the right interpretability–accuracy tradeoff, despite the original nature of fuzzy logic, has given rise to a growing interest in methods that take both aspects into account [7]–[11]. Ideally, both criteria should be satisfied to a high degree. However, since they are in conflict, this is not generally possible.

One way of doing this is to improve system's accuracy while trying to maintain interpretability to an acceptable level [9], [12]. By considering structural criteria, we can distinguish two main kinds of approaches that also take into account the interpretability of FRBSs.

1) *Complexity-based interpretability:* These approaches are used to decrease the complexity of the model that is obtained [12]–[21] (which are usually measured as the number of rules (NRs), variables, labels per rule, etc.).
2) *Semantics-based interpretability:* These approaches are used to preserve the semantics associated with the membership functions (MFs) [22]–[32]. We can find approaches that ensure semantic integrity, which usually

Manuscript received April 15, 2009; revised August 19, 2009 and November 13, 2009; accepted December 15, 2009. Date of publication January 19, 2010; date of current version May 25, 2010. This work was supported by the Spanish Ministry of Education and Science under Grant TIN2008-06681-C06-01.

M. J. Gacto is with the Department of Computer Science, University of Jaén, Jaén 23071, Spain (e-mail: mjgacto@ugr.es).

R. Alcalá and F. Herrera are with the Department of Computer Science and Artificial Intelligence, University of Granada, Granada 18071, Spain (e-mail: alcala@decsai.ugr.es; herrera@decsai.ugr.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TFUZZ.2010.2041008

imposes constraints on the MFs by considering measures such as distinguishability, coverage, fuzzy ordering, etc.

However, by paying attention to accuracy, one of the most widely used approaches to enhance the performance of FRBSs is the *tuning* of the MFs [27], [33]–[39]. It involves the improvement of a previous definition of the database (DB) once the RB has been obtained. The tuning methods refine the parameters that identify the MFs associated with the labels that comprise the DB [40]. Even though this approach is able to obtain highly accurate models, the semantic interpretability could be affected, depending on the variations that are performed in the MFs' shapes. The complexity of the models can also be a problem when a tuning is needed since usually, an excessive NRs is initially required to reach the highest degree of accuracy. Therefore, when an MF tuning is performed, three different criteria are required for a good accuracy–interpretability tradeoff: accuracy, complexity, and semantic interpretability.

A good way of optimizing these criteria simultaneously is the use of multiobjective evolutionary algorithms (MOEAs) [41], [42]. In fact, since this problem is multiobjective, most of the approaches that also take into account interpretability (especially, the complexity-based interpretability) use MOEAs to obtain a set of solutions with different degrees of accuracy and interpretability [13]–[15], [17]–[21], [23], [26].

In this paper, we propose an index to preserve the semantic interpretability of the DB while a tuning of the MFs is performed. The proposed index, i.e., GM3M, is defined as the geometric mean of three metrics, with the aim to minimize the displacement of the central point of the MFs, thus conserving the lateral amplitude rate of the MFs and maintaining the area of the original MFs that are associated with the linguistic labels. This measure can be used to quantify the interpretability of the tuned DB and could, therefore, be used as an objective within a multiobjective evolutionary process. To this end, we apply a specific MOEA to obtain interpretable and also accurate linguistic fuzzy models by concurrently performing a rule selection [16], [17], [43] and a tuning of the MF parameters with the following three objectives: *minimization of the system error*, *minimization of the NRs*, and *maximization of the proposed* Gm3m *index*. This postprocessing algorithm is based on the well-known *modified strength Pareto evolutionary algorithm* (*SPEA2*) [44]. It is called tuning and selection (TS) by SPEA2 for semantics-based index ($TS_{SP2\text{-}SI}$). In order to improve its ability to search, $TS_{SP2\text{-}SI}$ implements such concepts as incest prevention and restarting [45] and incorporates the main ideas of the algorithm proposed in [13] to guide the search toward the desired Pareto zone. Thus, $TS_{SP2\text{-}SI}$ aims to generate a complete set of Pareto-optimum solutions, with different tradeoffs between accuracy and interpretability in the double sense, thus decreasing the complexity and maintaining the semantic-based interpretability. We have not considered the well-known nondominated sorting genetic algorithm version II (NSGA-II) [46] since, in [13], approaches based on SPEA2 were shown to be more effective when a tuning of the MFs is performed.

We tested our approach on nine real-world regression datasets. In order to analyze the interaction between the fuzzy rule selection and the tuning of MFs and how it can affect the different objectives, these are also individually proved in a multiobjective framework and compared with their respective single-objective counterparts [35]. We compared the different approaches by applying nonparametric statistical tests for pairwise and multiple comparisons [47]–[50] by considering three representative points from the obtained Pareto fronts in the case of the MOEAs. Results confirm the effectiveness of our approach, and a wide range of solutions is obtained, which are not only more interpretable but also more accurate.

Section II briefly analyzes the state of the art on interpretable linguistic FRBS modeling. Section III introduces the rule-selection and the tuning techniques, which are used concurrently in this paper. Section IV presents the proposed index to control the semantic interpretability of the MFs. Section V presents the $TS_{SP2\text{-}SI}$ algorithm and describes its main characteristics, as well as the considered genetic operators. Section VI shows the experimental study and the results obtained. Finally, in Section VII, we point out some conclusions. An Appendix has been included to describe the nonparametric tests that are used in our study.

II. Introduction to Interpretability on Linguistic Modeling

This section reviews some basic ideas and works on the linguistic modeling interpretability. Along with the review in [10], which widely represents most of the existing works in the specialized literature, a framework to categorize fuzzy model interpretability into high-level interpretability and low-level interpretability has been recently suggested in [11].

1) High-level interpretability is obtained on the fuzzy rule level by conducting overall complexity reduction in terms of some criteria, such as a moderate number of variables, a moderate NRs, completeness, and consistency of rules (complexity-based interpretability).
2) Low-level interpretability of fuzzy models is achieved on fuzzy set level by optimizing MFs in terms of the semantic criteria on MFs (semantics-based interpretability).

The complexity-reduction techniques that are used in traditional system modeling can serve as fuzzy rule optimization, which corresponds to aiming at the parsimony of the fuzzy RB, which is one of the main high-level interpretability criteria of fuzzy systems. This clarification is helpful as there are plentiful traditional system modeling methods on complexity reduction that have great potentials to induce compact RB in fuzzy system modeling. Earlier works [16], [17] used rule selection on an initial set of classification rules and two different criteria: accuracy and NRs. Along with the work presented in [17], Ishibuchi and coworkers [18]–[20] optimized such complexity criteria by applying MOEAs. Rule length (which is, sometimes, used in combination with the NRs) has been included to minimize the length of the rules by either rule selection [14], [18], [19] or rule learning [18], [20], [21]. A method has also been proposed in [13] and deeply discussed in [15] to minimize the NRs along with a tuning of the MFs.

Low-level interpretability is achieved by optimizing MFs on the fuzzy set level. Specifically, low-level interpretability hails from the improvement on interpretability by introducing semantic constraint criteria into fuzzy modeling, which focus on the changes of MFs [11]. Classic approaches, such as [31] and [32], defined some helpful semantic criteria such as distinguishability, moderate number of MFs, natural zero positioning, normality, and coverage. These properties were later included in an MOEA to check their interaction when they evolve simultaneously [26]. Other works have focused on defining proper similarity metrics as a way to measure the distinguishability and coverage of the MFs [28], which are sometimes used to fix some minimum values of covering [24], [27], and some others are used to define maximum values of similarity for merging fuzzy sets and rules (particularly when MFs came from clustering techniques) [25], [30]. A similarity measure is also optimized in [29] to promote a good covering of the MFs, along with two complexity criteria in a combined index. Another MOEA is adopted in [23] to perform context adaptation. This algorithm considers the system error and an interpretability index to preserve the fuzzy ordering and a good distinguishability.

Additionally, some other works try to go a step ahead by considering all these kinds of measures in a linguistic framework in order to search for a more global definition of interpretability [12], [22]. In this sense, a conceptual framework is presented in [7] to characterize the interpretability of FRBSs. It makes reference to [10] and [11], which are combined in several interpretability levels (extending the low–high categorization).

Although most of the sematic-based approaches are mainly focused on finding partitions with a good overlapping among MFs (covering and distinguishability), in this paper, since interpretability is dependent on the problem context and user perceptions, we try to keep partitions and meanings to their original values, while performance improvements are still allowed. Further, it has also been combined with one of the classic complexity measures.

III. Fuzzy Rule Selection and Tuning of Membership Functions

In this paper, we present an MOEA for postprocessing that concurrently performs a fuzzy rule selection and a tuning of the MFs. This section briefly introduces the fuzzy-rule-selection technique and the tuning approach used to optimize the MF parameters.

A. Fuzzy Rule Selection

Fuzzy-rule-set-reduction techniques try to minimize the NRs of a given FRBS while maintaining (or even improving) the system's performance. To do this, erroneous and conflicting rules that degrade the performance are eliminated, thus obtaining a more cooperative fuzzy rule set and, as a result, potentially improving system's accuracy. Furthermore, in many cases, accuracy is not the only requirement of the model, but interpretability also becomes an important aspect. Reduction of the model complexity is a way to improve the system's readability, i.e., a compact system with few rules generally requires less effort in interpretation. Fuzzy-rule-set-reduction techniques are usually applied as a postprocessing stage once an initial fuzzy

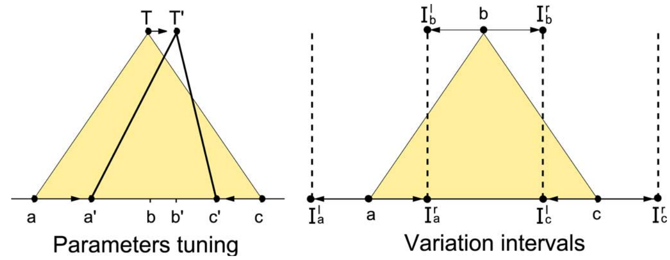


Fig. 1. Tuning by changing the basic MF parameters and the variation intervals.

rule set has been extracted. One of the most used fuzzy-rule-set-reduction techniques is the rule selection. This approach involves obtaining an optimal subset of fuzzy rules from a previous fuzzy rule set by selecting some of them. We may find several methods for rule selection, with different search algorithms that look for the most successful combination of fuzzy rules [16], [17], [43]. An interesting heuristic rule-selection procedure is proposed in [51], where, by means of statistical measures, a relevance factor is computed for each fuzzy rule in the FRBSs to subsequently select the most relevant ones.

These kinds of techniques for rule selection could be easily combined with other postprocessing techniques to obtain more compact and accurate FRBSs. This way, some works have considered the selection of rules along with the tuning of MFs by coding all of them (rules and parameters) in the same chromosome [13], [15], [33]–[35] within the same process and considering only performance criteria. Rules would be extracted only if it is possible to either maintain or even improve the system's accuracy. A very interesting conclusion from some of these recent works [15], [35] is that both techniques can present a positive synergy when they are combined within a well-designed optimization process.

B. Tuning of Membership Functions

This approach, which is usually called DB tuning, involves refining the MF shapes from a previous definition once the remaining FRBS components have been obtained [27], [36]–[39]. The classic way to refine the MFs is to change their definition parameters. For example, if the following triangular-shaped MF is considered:

$$\mu(x) = \begin{cases} \frac{x-a}{b-a}, & \text{if } a \le x < b \\ \frac{c-x}{c-b}, & \text{if } b \le x \le c \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

changing the basic parameters—a, b, and c—will vary the shape of the fuzzy set that is associated with the MF, thus influencing the FRBS performance (see Fig. 1). This is also true for other shapes of MFs (trapezoidal, Gaussian, etc.).

Tuning involves fitting the characterization of the MFs associated with the primary linguistic terms that are considered in the system. Thus, the meaning of the linguistic terms is changed from a previous definition (i.e., an initial DB that is composed of the sematic concepts, and the corresponding MFs give meaning to them). As said, in order to preserve the semantic integrity throughout the MF-optimization process [9], [31], [32], some researchers have proposed several properties. Considering one or more of these properties, several semantic constraints can be applied in the design process in order to obtain a DB that maintains the linguistic model integrity to the highest possible level [22], [24], [25], [29], [36].

In this paper, in order to illustrate the performance of the proposed approach, we use equidistributed strong fuzzy partitions [52] to define an initial set of triangular MFs. These kinds of fuzzy partitions, in which the sum of the membership degrees within the variable domain are equal to 1.0 and the triangular MFs are equidistant (therefore, symmetrical), perfectly meet the required semantic constraints, and they are widely assumed to have a high level of transparency. Anyhow, the initial DB should be given by an expert, if possible, since the concepts and their meaning strongly depend on the problem and the person who makes the assessment. In order to maintain the semantic integrity, we also consider some basic constraints by defining convenient variation intervals for each MF parameter. For each $\text{MF}_j = (a_j, b_j, c_j)$, where $j = (1, \ldots, m)$, and m is the number of MFs in a given DB, the variation intervals are calculated as follows (see Fig. 1):

$$\begin{aligned} [I^l_{a_j}, I^r_{a_j}] &= [a_j - (b_j - a_j)/2, a_j + (b_j - a_j)/2] \\ [I^l_{b_j}, I^r_{b_j}] &= [b_j - (b_j - a_j)/2, b_j + (c_j - b_j)/2] \\ [I^l_{c_j}, I^r_{c_j}] &= [c_j - (c_j - b_j)/2, c_j + (c_j - b_j)/2]. \end{aligned} \tag{2}$$

Due to these restrictions, it is possible to maintain the integrity of MFs to a reasonable level. In any case, it would be very interesting to have a measure for the quality of the tuned MFs. We propose three metrics that try to preserve the original form of the MFs, thus improving, if possible, the tradeoff between accuracy and interpretability.

IV. Semantic-Based Interpretability Index

In this section, we propose several metrics to measure the interpretability when a tuning is performed on the DB. At this point, we should remark that these metrics are based on the existence of the variation intervals (integrity constraints) that are defined in the previous section and, therefore, on the assumption that the initial DB comprises triangular MFs. Even though these measures and index are proposed to work with triangular MFs, they can be easily extended with some small changes in the formulation of Gaussian or trapezoidal MFs. Since significant changes in the DB can have a negative influence on interpretability, each metric is proposed to control how good some desirable aspects of the tuned MFs are with respect to the original ones (relative, not absolute, metrics). The metrics proposed are the following.

1) MFs displacement (δ): This metric measures the proximity of the central points of the MFs to the original ones. The closer they are to the original points, the higher the displacement.
2) MFs lateral amplitude rate (γ): This metric measures the left/right rate differences of the tuned and the original MFs. The closer the rates are, the higher the lateral amplitude rate.
3) MFs area similarity (ρ): This metric measures the area similarity of the tuned and the original MFs. It should be higher if the tuned and the original areas are closer.

In the following sections, the three proposed metrics will be explained in depth.

A. MFs Displacement Measure (δ)

This metric can control the displacements in the central point of the MFs. It is based on computation of the normalized distance between the central points of the tuned MF and the original MF, and is calculated by obtaining the maximum displacement attained on all the MFs. For each MF_j in the DB, we define $\delta_j = |b_j - b'_j|/I$, where $I = (I^r_{b_j} - I^l_{b_j})/2$ represents the maximum variation for each central parameter. Thus, δ^* is defined as $\delta^* = \max_j\{\delta_j\}$. The δ^* metric takes values between 0 and 1; therefore, values near 1 show that the MFs present a great displacement. The following transformation is made so that this metric represents proximity (maximization):

$$\text{Maximize } \delta = 1 - \delta^*. \tag{3}$$

This metric could also be used for either Gaussian or trapezoidal MFs by considering the middle of the core as the position to preserve.

B. MFs Lateral Amplitude Rate Measure (γ)

This metric can be used to control the shapes of the MFs. It is based on relating the left and right parts of the support of the original and the tuned MFs. Let us define left $S_j = |a_j - b_j|$ as the amplitude of the left part of the original MF support and right $S_j = |b_j - c_j|$ as the right-part amplitude. Let us define left $S'_j = |a'_j - b'_j|$ and right $S'_j = |b'_j - c'_j|$ as the corresponding parts in the tuned MFs. The variable γ_j is calculated using the following equation for each MF:

$$\gamma_j = \frac{\min\{\text{left } S_j/\text{right } S_j, \text{left } S'_j/\text{right } S'_j\}}{\max\{\text{left } S_j/\text{right } S_j, \text{left } S'_j/\text{right } S'_j\}}. \tag{4}$$

Values near 1 mean that the left and right rates in the original MFs are highly maintained in the tuned MFs. Finally, γ is calculated by obtaining the minimum value of γ_j as

$$\text{Maximize } \gamma = \min_j\{\gamma_j\}. \tag{5}$$

This metric always presents a value of 1 in the case of Gaussian MFs. It could also be used for trapezoidal MFs by considering the middle of the core as the central point, computing γ_j with the core extremes, computing γ_j with the MF extremes, and averaging both values.

C. MFs Area Similarity Measure (ρ)

This metric can be used to control the area of the shapes of the MFs. It is based on relating the areas of the original and the tuned MFs. Let us define A_j as the area of the triangle that represents the original MF_j and A'_j as the new area. The variable

ρ_j is calculated using the following equation for each MF:

$$\rho_j = \frac{\min\{A_j, A'_j\}}{\max\{A_j, A'_j\}}. \quad (6)$$

Values near 1 mean that the original area and the tuned area of the MFs are more similar (fewer changes). The ρ metric is calculated by obtaining the minimum value of ρ_j

$$\text{Maximize } \rho = \min_j\{\rho_j\}. \quad (7)$$

This metric is also applicable for trapezoidal and Gaussian MFs.

D. Semantics-Based Interpretability Index Based on Aggregation of the Three Measures: GM3M

We propose an aggregation of the metrics in a global index based on the geometric mean. As mentioned, this index is called GM3M and is defined as

$$\text{Maximize GM3M} = \sqrt[3]{\delta\gamma\rho}. \quad (8)$$

The value of GM3M ranges between 0 (which is the lowest level of interpretability) and 1 (which is the highest level of interpretability). The use of either $\min_j\{\cdot\}$ or $\max_j\{\cdot\}$ to compute the different metrics ensures the interpretability to a minimum level in all the MFs, since our main aim is to measure the worst case. Therefore, if there is a major problem in any of the MFs, it can be detected and reflected in each particular metric. Similarly, it is clear that if only one of the metrics has very low values, a problem arises in the interpretability. The used aggregation operator considers this fact. Moreover, all these relative metrics present complementary properties to measure the relation with the initial MFs.

V. Multiobjective Evolutionary Algorithm for Rule Selection and Tuning of Fuzzy Rule-Based Systems

Since it is not possible to either obtain the different interpretability–accuracy tradeoff degrees or handle the synergy of both approaches separately, the proposed algorithm performs a fuzzy rule selection along with a tuning of the MFs in order to improve the system's accuracy as a first objective, the model complexity as a second objective, and the GM3M index in order to preserve the semantic interpretability as the third objective. As mentioned, it is a specific MOEA that is called SPEA2 for semantic interpretability, i.e., $\text{TS}_{\text{SP2-SI}}$, which is based on the well-known SPEA2 [44] algorithm. In the next section, the main components of this algorithm are described, and then, the specific characteristics and its main steps are presented.

A. Objectives

Every chromosome is associated with a 3-D objective vector, each element of which expresses the fulfillment degree of the following three objectives:

1) semantic interpretability maximization: semantic-based index, GM3M;
2) complexity minimization: number of selected rules, NR;
3) error minimization: mean-squared error divided by 2 ($\text{MSE}_{/2}$).

The number of input variables is another complexity measure that could be considered to improve the system's interpretability. However, we have not used this measure since this can be considered in a previous stage, thus avoiding the use of a fourth objective in the MOEAs, which, nowadays, are not able to work properly with such quantity of objectives. The value of $\text{MSE}_{/2}$ of an FRBS that is decoded from a given chromosome is defined as follows: $\text{MSE}_{/2} = (1/2)|D|\sum_{l=1}^{|D|}(F(x^l) - y^l)^2$, where $|D|$ is the dataset size, $F(x^l)$ is the output of the FRBS when the lth example is an input, and y^l is the known desired output. The fuzzy inference system uses the *center of gravity weighted by the matching* strategy as a defuzzification operator and the *minimum t-norm* as implication and conjunctive operators.

B. Coding Scheme and Initial Gene Pool

A double coding scheme for both *rule selection* (C_{S}) and *tuning* (C_{T}) is used: $C^p = C_{\text{S}}^p C_{\text{T}}^p$. In the $C_{\text{S}}^p = (c_{\text{S}1}, \ldots, c_{\text{S}m})$ part, the coding scheme consists of binary-coded strings with size m (where m is the number of initial rules). Depending on whether a rule is selected or not, values of either "1" or "0" are, respectively, assigned to the corresponding gene. In the C_{T} part, a real coding is used, with m^i being the number of labels of each of the n variables in the DB

$$C_{\text{T}}^p = C_1 C_2 \ldots C_n$$
$$C_i = (a_1^i, b_1^i, c_1^i, \ldots, a_{m^i}^i, b_{m^i}^i, c_{m^i}^i), \qquad i = 1, \ldots, n.$$

The initial population is obtained with all individuals having all genes with value "1" in C_{S}. In the C_{T} part, the initial DB is included as a first individual, and the remaining individuals are generated at random within the corresponding variation intervals that are defined in Section III-B.

C. Crossover and Mutation

In this section, we propose an intelligent crossover and a mutation operator based on our experience in this problem. This is able to adequately profit from the parents when both rule selection and tuning are applied. The steps to obtain each offspring are as follows.

1) Blend crossover (BLX)-0.5 [53] is applied to obtain the C_{T} part of the offspring.
2) Once the offspring C_{T} part has been obtained, the binary part C_{S} is attained based on the C_{T} parts (MFs) of parents and offspring. For each gene in the C_{S} part that represents a concrete rule, the following hold.
    a) The MFs involved in such rule are extracted from the corresponding C_{T} parts for each individual that is involved in the crossover (offspring and parents 1 and 2). Thus, we can obtain the specific rules that each of the three individuals represent.
    b) Euclidean normalized distances are computed between the offspring rule and each parent rule by considering the center points (vertex) of the MFs that are composed of such rules. The differences

between each pair of centers are normalized by the amplitudes of their respective variation interval.

c) The parent with the rule closer to the one that is obtained by the offspring is the one that determines whether this rule is selected or not for the offspring by directly copying its value in C_{S} for the corresponding gene.

This process is repeated until all the C_{S} values are assigned for the offspring. Four offspring are obtained by repeating this process four times. (After considering mutation, only the two most accurate values are taken as descendants.) By applying this operator, exploration is performed in the C_{T} part, and C_{S} is directly obtained based on the previous knowledge that each parent has about the fact whether a specific configuration of MFs can be used for each rule. This avoids the possibility of recovering a bad rule that was discarded for a concrete configuration of MFs, while allowing the recovery of a good rule that is still considered for this concrete configuration, thus increasing the probability of success in either the selection or the elimination of a rule for each concrete configuration of MFs. Since a better exploration is performed for the C_{S} part, the mutation operator does not need to add rules. This way, once an offspring is generated, the mutation operator changes a gene value at random in the C_{T} part and directly sets to zero a gene that is selected at random in the C_{S} part (one gene is modified in each part) with probability P_m.

By applying these operators, two problems are solved. First, crossing individuals with very different rule configurations is more productive. Second, this way of working favors rule extraction since mutation is employed only to remove unnecessary rules.

D. Main Characteristics of $TS_{SP2\text{-}SI}$

The proposed algorithm uses the SPEA2-selection mechanism. However, in order to improve the algorithm's ability to search, the following changes are considered.

1) The proposed algorithm includes a mechanism for incest prevention based on the concepts of CHC [45] in order to avoid premature convergence in the C_{T} part (real coding), which is the main responsibility of accuracy improvements and represents a more complicated search space than the C_{S} part (binary coding). In CHC, only those parents are crossed whose Hamming distance divided by 4 is greater than a threshold. Since we consider a real coding scheme (i.e., only C_{T} parts are considered), we have to transform each gene using a gray code with a fixed number of bits per gene (BGene), which are determined by the system's expert. This way, the threshold value is initialized as $L = (\#C_{\mathrm{T}} \times \text{BGene})/4$, where $\#C_{\mathrm{T}}$ is the number of genes in the C_{T} part of the chromosome. At each generation of the algorithm, the threshold value decreases by 1, which allows crossing closer solutions. This mechanism can also be maintained because the parent selection is multiobjective, which provides a parent diversity that is similar to the original CHC.
2) The restarting operator forces the external population to be empty and generates a new initial population. This initial population includes a copy of the individuals with the best value in each objective (before removing them from the external population). The remaining individuals in the new population take the values of the most accurate individual in the C_{S} part and values generated at random in the C_{T} part. This preserves the most accurate and the most interpretable solutions that are obtained. The restarting operator is applied when we detect that all the crossovers are allowed. However, in order to avoid premature convergence, we apply the first restart if 50% of crossovers are detected at any generation (the required ratio can be defined as $\%_{\text{required}} = 0.5$). This condition is updated each time restarting is performed as $\%_{\text{required}} = (1+\%_{\text{required}})/2$. Moreover, the most accurate solution should be improved before each restart. To preserve a well-formed Pareto front, the restart is not applied at the end. The number of evaluations without restart can be estimated as the number of evaluations needed to apply the first restart multiplied by 10. Additionally, restart is disabled if it was never applied before reaching the midpoint of the total number of evaluations.
3) At each stage of the algorithm (between restarting points), the number of solutions in the external population ($\overline{P}_{t+1}$) that is considered to form the mating pool is progressively reduced, by focusing only on those with the best accuracy. To do this, the solutions are sorted from the best to the worst (considering accuracy as criterion), and the number of solutions that are considered for selection is reduced progressively from 100% at the beginning to 50% at the end of each stage. It is done by taking into account the value of L. In the last evaluations when restart is disabled, the mechanism to focus on the most accurate solutions (which is the most difficult objective) is also disabled to obtain a wide, well-formed Pareto front, from the most accurate solutions to the most interpretable ones.

The main steps of $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ are finally presented in Fig. 2 (see SPEA2 in [44]).

VI. Experimental Study

To evaluate the usefulness of the proposed approach, we used nine real-world problems. Table I summarizes the main characteristics of the nine datasets and shows the link to the knowledge extraction based on evolutionary learning (KEEL) software tool Web page (http://www.keel.es/) [55] from which these can be downloaded. This section is organized as follows.

1) Section VI-A presents the experimental setup.
2) Section VI-B analyzes the tuning of MFs individually, by paying attention to the GM3M index. To this end, the tuning component of the proposed approach and its single-objective counterpart are compared in terms of the most accurate solutions. Some example DBs are also presented in order to graphically show the effects of the use of the GM3M index as an objective in the evolutionary model.
3) Section VI-C presents an analysis on the rule selection individually. In order to better analyze the interaction between the different components of the proposed approach,

Input: N (population size), $\overline{N}$ (external population size), E (maximum number of evaluations),
$BGene$ (bit per gene for gray code).
Output: A (non-dominated set).
Terminology:
$\#C_T$ (number of genes in the real part C_T), $\#O$ (number of objectives), Evs (current number of evaluations),
L (threshold for incest prevention), $InitL = (\#C_T * BGene)/4$ (initial threshold), $R\%$ (descendant % required for restart), Rst (variable to activate restart), $Nded$ (evaluations needed to form a Pareto),
Acc^{+} (accuracy improvement is detected in the most accurate solution from the latest restart).

Algorithm:
1) Generate P_0 (initial population) and create $\overline{P}_0 = \emptyset$ (empty external population).
2) Evaluate individuals in P_0 ($MSE_{/2}$) and set:
   – $L = InitL$; $R\% = 0.5$; $Rst = false$; $Evs = N$; $Nded = 0$; $t = 0$;
3) Calculate fitness values of individuals in P_t and $\overline{P}_t$. Copy all non-dominated individuals in $P_t \cup \overline{P}_t$ to $\overline{P}_{t+1}$. If $|\overline{P}_{t+1}| > \overline{N}$ apply truncation operator. If $|\overline{P}_{t+1}| < \overline{N}$ fill with dominated in $P_t \cup \overline{P}_t$.
4) If $Evs \geq E$, return A and stop.
5) If (Rst) and $(Evs < E - Nded)$ and (Acc^{+}):
   – $L = InitL$; $R\% = (R\% + 1)/2.0$; If $Nded$ is 0, $Nded = Evs * 10$; $Rst = false$.
   – Copy the best individuals in each objective to P_t. Empty $\overline{P}_t$ ($\overline{P}_t = \emptyset$). Fill remaining $N - \#O$ individuals in P_t with C_T at random and C_S equal to the most accurate individual.
   – Evaluate the $N - \#O$ new individuals in P_t ($MSE_{/2}$), set $Evs += N - \#O$ and go to Step 3.
6) Generate the next population:
   – If $Evs < E - Nded$, set $P = (L/(InitL * 2.0) + 0.5)$ else set $P = 1.0$. Perform binary tournament selection with replacement on the $\lfloor \overline{N} * P \rfloor$ most accurate solutions of $\overline{P}_{t+1}$ to fill the mating pool.
   – Apply crossover (BLX-Specific) and mutation for each two parents in the mating pool if the hamming distance between their C_T part Gray codings divided by 4 is over L.
   – Set P_{t+1} to the resulting population with the obtained G descendant. Set $Evs += G * 2$.
7) Variables updating:
   – If $L > 0$, $L = L - 1$; If $G \geq N * R\%$, $Rst = true$; If $Nded$ is 0 and $Evs \geq E/2$, $Nded = E$.
8) Go to Step 3 with $t = t + 1$.

Fig. 2. $\text{TS}_{\text{SP2-SI}}$ algorithm scheme.

the rule-selection component has also been compared with its single-objective counterpart.

4) Section VI-D analyzes the proposed approach and the interaction between the tuning and the rule-selection components. This analysis has been carried out in the same way, i.e., by considering only tuning and paying attention to the effects that the concurrent use of both techniques promotes to the different criteria, particularly to the GM3M index.
5) Section VI-E includes a global statistical analysis of the most accurate solutions by considering all the approaches and the corresponding optimized measures/objectives.
6) Finally, Section VI-F shows a graphical and statistical analysis of the obtained Pareto fronts. To perform this study, we plot the centroids (average values) of three representative points of the Pareto fronts (from the most accurate to the most interpretable) on the accuracy–complexity and accuracy–semantic planes. These plots provide a glimpse of the trend of the Pareto fronts. We also present a statistical analysis of the centroids of the most interpretable and intermediate solutions. For completeness, we also show some representative Pareto fronts that are achieved by the different MOEAs.

TABLE I
DATASETS THAT ARE CONSIDERED FOR THE EXPERIMENTAL STUDY

Datasets	Name	Variables	Patterns
Plastic Strength	PLA	3	1650
Quake	QUA	4	2178
Electrical Maintenance	ELE	5	1056
Abalone	ABA	9	4177
Stock prices	STP	10	950
Weather Ankara	WAN	10	1609
Weather Izmir	WIZ	10	1461
Mortgage	MOR	16	1049
Treasury	TRE	16	1049

Available at: http://sci2s.ugr.es/keel/datasets.php

A. Experimental Setup

In all the cases, the well-known *ad hoc* data-driven learning algorithm of Wang and Mendel [54] is applied to obtain an initial set of candidate linguistic rules. The initial linguistic partitions comprise *five linguistic terms* in the case of datasets with less than nine variables and *three linguistic terms* in the remaining ones (which helps obtain a more reasonable NRs in the more complex datasets). Once the initial RB is generated, the different postprocessing algorithms can be applied. The methods that are considered for the experiments are briefly described in Table II. In order to evaluate the advantages of concurrently performing rule selection and tuning for the optimization of the three objectives simultaneously ($\text{TS}_{\text{SP2-SI}}$), we also analyze the use of the multiobjective approach in both rule selection and tuning separately. In practice, we consider chromosomes that are composed of only the C_{S} part for the rule selection (S_{SP2}) and the C_{T} part for the tuning of MFs ($\text{T}_{\text{SP2-SI}}$). Further, their single-objective accuracy-oriented counterparts are also considered in order to analyze the influence of the interpretability criteria in the most difficult one (accuracy).

Clearly, it would make no sense to consider either the GM3M objective when no tuning is performed or the NR objective when no rule selection is performed. It is assumed that the approaches that perform only rule selection have the maximum semantic interpretability and those that perform tuning have the worst NR. Accordingly, the approaches that consider only rule selection should be compared in the accuracy–complexity (MSE–NR) plane, while the approaches that consider only tuning should be compared in the accuracy–semantic (MSE–GM3M) plane. In the case of the proposed method, i.e., $\text{TS}_{\text{SP2-SI}}$, which uses the three objectives, we project the solutions that are obtained in both planes, accuracy–complexity and accuracy–semantic, subsequently removing the dominated solutions that appear from these projections. This way, the methods that perform rule selection and tuning concurrently can be compared with the methods that perform only rule selection in the accuracy–complexity plane and with those that perform only tuning in the accuracy–semantic plane. Some researchers have also used these kinds of

TABLE II
METHODS THAT ARE CONSIDERED FOR COMPARISON WITH CLASSICAL TUNING

Method	Ref.	Description	Objectives
WM	[54]	Wang & Mendel Algorithm (Initial RB Generation)	—
		Single-Objective Methods for Post-processing	
S	[35]	Genetic Rule Selection	$MSE_{/2}$
T	[35]	Genetic Tuning of Parameters	$MSE_{/2}$
TS	[35]	Genetic Tuning and Rule Selection	$MSE_{/2}$
		Multi-Objective Evolutionary Algorithms for Post-processing	
$\mathbf{S}_{SP2}$	—	Rule Selection by SPEA2	$MSE_{/2}$ / NR
$\mathbf{T}_{SP2-SI}$	—	Tuning with semantic by SPEA2	$MSE_{/2}$ / GM3M
$\mathbf{TS}_{SP2-SI}$	Proposed here	Tuning and Rule Selection with semantic by SPEA2	$MSE_{/2}$ / NR / GM3M

projections for graphical representation when three objectives are optimized simultaneously [18].

In all the experiments, we adopted a *fivefold cross-validation model*, i.e., we randomly split the dataset into five folds, each containing 20% of the patterns of the dataset, and used four folds for training and one for testing.[1] For each of the possible five different partitions (training/test), the algorithm was applied six times, considering a different seed for the random-number generator. Therefore, we consider the average results of 30 runs. In the case of methods with a multiobjective approach, for each dataset and for each trial, we generate the approximated Pareto front in the corresponding objective planes. Then, we focus on three representative points: the most interpretable (MAX INT), the median (MEDIAN INT/ACC), and the most accurate in training (MAX ACC) points. For each representative point, we compute the mean values over the 30 trials of the MSEs on the training and test sets (i.e., $\mathrm{MSE}_{/2}^{\mathrm{tra}}$ and $\mathrm{MSE}_{/2}^{\mathrm{tst}}$), the NR, and/or the GM3M index, depending on the objective planes that are involved. For the single-objective-based approaches, we compute the same mean values over the 30 solutions that are obtained for each dataset. These three points are representative positions on each plane, i.e., accuracy–complexity or accuracy–semantic, and they have been considered only to perform a statistical analysis on the different planes. Besides, the final user could select the most appropriate solution from the final Pareto front by also looking for a tradeoff between NR and GM3M, depending on its own preferences.

In order to assess whether significant differences exist among the results, we adopt statistical analysis [47]–[50] and, in particular, nonparametric tests, according to the recommendations made in [47] and [48], where a set of simple, safe, and robust nonparametric tests for statistical comparisons of classifiers has been introduced. For pairwise comparison, we use Wilcoxon's signed-ranks test [56], [57], and for multiple comparisons, we will employ different approaches, including Friedman's test [58], Iman and Davenport's test [59], and Holm's method [60]. A detailed description of these tests is presented in Appendix. To perform the tests, we use a level of confidence $\alpha = 0.1$. In particular, Wilcoxon's test is based on computing the differences between two sample means (typically, mean test errors that are obtained by a pair of different algorithms on different datasets). In the classification framework, these differences are well defined since these errors are in the same domain. In our case, to have well-defined differences in $\mathrm{MSE}_{/2}$ and NR (it is not necessary in the case of GM3M), we propose to adopt a normalized difference DIFF, which is defined as

$$\mathrm{DIFF} = \frac{\mathrm{Mean(Other)} - \mathrm{Mean(Reference\ Algorithm)}}{\mathrm{Mean(Other)}} \quad (9)$$

where $\mathrm{Mean}(x)$ represents either the $\mathrm{MSE}_{/2}$ or the NR means that are obtained by the x algorithm. This difference expresses the improvement in percentage of the reference algorithm.

TABLE III
INITIAL RESULTS OBTAINED BY WM

Dataset	NR	$MSE_{/2}^{tra}$	$MSE_{/2}^{tst}$	σ_{tra}	σ_{tst}
PLA	14.8	3.434	3.557	0.265	0.235
QUA	53.6	0.0258	0.0267	0.001	0.002
ELE	65.0	57606	57934	2841	4733
ABA	68	8.407	8.422	0.443	0.545
STP	122.8	9.074	9.042	0.486	0.809
WAN	156.0	16.063	16.393	0.961	1.700
WIZ	104.8	6.944	7.368	0.720	0.909
MOR	77.6	0.985	0.973	0.129	0.090
TRE	75.0	1.636	1.631	0.121	0.181

The average results of the initial FRBSs, along with their standard deviations (reference results), which are obtained by WM in the five folds, are shown in Table III. In the case of the studied postprocessing algorithms, the values of the input parameters that are considered by the single-objective methods are as follows: population size of 61, 100 000 evaluations, 0.6 as crossover probability, and 0.2 as mutation probability per chromosome. In the case of the MOEAs, these are the following: population size of 200, external population size of 61, 100 000 evaluations, 0.2 as mutation probability, and 30 bits per gene for the Gray codification.

B. Analysis on the Tuning of MFs and the Semantic-Based Index: GM3M

This section analyzes the performance of the methods that perform only tuning of the MFs. Table IV shows the results obtained by T and the results obtained by $\mathrm{T}_{\mathrm{SP2\text{-}SI}}$ in the three representative points of the accuracy–semantic plane, which are used further for a statistical analysis of the

[1]The corresponding data partitions (fivefold) for these datasets are available at the KEEL project Web page [55]: http://sci2s.ugr.es/keel/datasets.php.

TABLE IV
RESULTS OBTAINED BY THE METHODS THAT PERFORM ONLY TUNING OF MFs

Dataset	Method	MAX INT						MEDIAN (INT/ACC)						MAX ACC					
		$MSE_{/2}^{tra}$	$MSE_{/2}^{tst}$	GM3M	(δ	γ	ρ)	$MSE_{/2}^{tra}$	$MSE_{/2}^{tst}$	GM3M	(δ	γ	ρ)	$MSE_{/2}^{tra}$	$MSE_{/2}^{tst}$	GM3M	(δ	γ	ρ)
PLA	T		-						-					1.200	1.251	0.259	(0.11	0.30	0.69)
	T_{SP2-SI}	3.349	3.474	0.985	(0.99	0.98	0.99)	1.572	1.614	0.780	(0.79	0.76	0.80)	1.194	**1.242**	**0.363**	(0.23	0.44	0.72)
QUA	T		-						-					0.0175	**0.0183**	0.097	(0.02	0.12	0.61)
	T_{SP2-SI}	0.0251	0.0260	0.953	(0.96	0.94	0.96)	0.0185	0.0194	0.607	(0.52	0.57	0.75)	0.0175	**0.0183**	**0.109**	(0.01	0.30	0.63)
ELE	T		-						-					17020	21027	0.225	(0.06	0.34	0.69)
	T_{SP2-SI}	56529	56983	0.988	(0.99	0.98	0.99)	24671	26870	0.735	(0.72	0.70	0.79)	15884	**19257**	**0.319**	(0.22	0.40	0.65)
ABA	T		-						-					2.688	2.770	0.144	(0.03	0.18	0.64)
	T_{SP2-SI}	7.886	7.897	0.965	(0.97	0.95	0.97)	3.563	3.611	0.801	(0.83	0.72	0.86)	2.648	**2.744**	**0.298**	(0.18	0.39	0.62)
STP	T		-						-					0.904	1.072	**0.095**	(0.02	0.08	0.60)
	T_{SP2-SI}	8.867	8.834	0.975	(0.98	0.96	0.98)	2.356	2.450	0.706	(0.73	0.65	0.76)	0.797	**0.921**	0.090	(0.01	0.21	0.60)
WAN	T		-						-					1.928	**2.287**	0.144	(0.03	0.19	0.68)
	T_{SP2-SI}	14.683	15.537	0.947	(0.95	0.93	0.95)	3.847	4.650	0.707	(0.69	0.68	0.77)	1.693	2.566	**0.240**	(0.11	0.39	0.68)
WIZ	T		-						-					1.103	1.561	0.155	(0.05	0.14	0.67)
	T_{SP2-SI}	6.516	6.729	0.949	(0.95	0.93	0.96)	1.892	2.037	0.697	(0.65	0.69	0.77)	1.048	**1.243**	**0.260**	(0.10	0.41	0.66)
MOR	T		-						-					0.050	0.061	0.168	(0.04	0.20	0.66)
	T_{SP2-SI}	0.947	0.935	0.975	(0.98	0.97	0.98)	0.245	0.254	0.749	(0.77	0.67	0.82)	0.036	**0.043**	**0.183**	(0.06	0.17	0.72)
TRE	T		-						-					0.052	0.063	**0.152**	(0.04	0.16	0.68)
	T_{SP2-SI}	1.527	1.522	0.951	(0.95	0.94	0.97)	0.303	0.310	0.730	(0.75	0.65	0.82)	0.047	**0.061**	0.113	(0.02	0.20	0.70)

TABLE V
WILCOXON'S TEST: T (R^+) VERSUS $\mathrm{T_{SP2\text{-}SI}}$ (R^-) ON GM3M AND $\mathrm{MSE}_{/2}^{\mathrm{tst}}$ AT MAX ACC

Comparison	Measure	R^+	R^-	Hypothesis ($\alpha = 0.1$)	p-value
T vs. T_{SP2-SI}	GM3M	5	40	Rejected	0.038
T vs. T_{SP2-SI}	$MSE_{/2}^{tst}$	6	39	Rejected	0.051

multiobjective methods. In addition to the semantic-based index, i.e., GM3M, we show the mean values of the three measures that comprise the index, δ, γ, and ρ. [In any event, we should take into account that $(\sum_{i=1}^{30} \mathrm{GM3M}_i/30) \neq \sqrt[3]{(\sum_{i=1}^{30} \delta_i/30)(\sum_{i=1}^{30} \gamma_i/30)(\sum_{i=1}^{30} \rho_i/30)}$.]

Table V shows the results of the Wilcoxon test on the test error and the GM3M measures for T and $\mathrm{T_{SP2\text{-}SI}}$ at MAX ACC. The results show that $\mathrm{T_{SP2\text{-}SI}}$ outperforms T on the test error and GM3M. The null hypothesis that is associated with Wilcoxon's test is rejected ($p < \alpha$) in both cases in favor of $\mathrm{T_{SP2\text{-}SI}}$ due to the differences between R^+ and R^-. This is due to the complex search space that the parametric tuning of MFs involves. The use of both objectives and the modified SPEA2 algorithm helps improve the exploration/exploitation tradeoff to find more optimal solutions.

Fig. 3 shows a representative example in ELE (same data partition and seed) of a DB that is obtained with T and three DBs that are obtained with $\mathrm{T_{SP2\text{-}SI}}$, with the first one with the most interpretable solution, the second one with the median solution, and the last one with the most accurate solution. The DBs obtained are shown in black and the initial DB is shown in gray. To ease graphic representation, the MFs are labeled from "l1" to "l5." Nevertheless, such MFs are associated with a linguistic meaning that is determined by an expert. With these examples, we show the expected correlation between the GM3M index and the semantic interpretability of the obtained DBs. It is quite interesting that the solution with the highest interpretability obtains about a 37% improvement in test with respect to WM and a value of GM3M near 1.

C. Analysis of the Rule Selection

In this section, we present a brief study on the methods that perform only rule selection. Table VI shows the results that are obtained by S and the results that are obtained by $\mathrm{S_{SP2}}$ in the three representative points of the accuracy–complexity plane.

In order to assess whether we can conclude that $\mathrm{S_{SP2}}$ statistically outperforms S in terms of test error and NR measure, we apply Wilcoxon's test to the results achieved by these algorithms in the most accurate solutions. Table VII shows the results of the application of Wilcoxon's test on these measures. The null hypothesis that is associated with the Wilcoxon's test is now accepted ($p > \alpha$) in both cases. Thus, we can conclude that the results achieved by S and $\mathrm{S_{SP2}}$ are statistically different neither on the test error nor on the NR measure. In this case, the search space is well handled by both approaches since equivalent results are obtained by considering the most accurate solutions of the obtained Pareto fronts. In any event, $\mathrm{S_{SP2}}$ is able to obtain a set of valid solutions with different accuracy–complexity tradeoffs.

D. Analysis of the Interaction of the Tuning With Rule Selection

This section analyzes the results of the proposed method, i.e., $\mathrm{TS_{SP2\text{-}SI}}$, which performs both rule selection and tuning of the MFs simultaneously, with respect to its single-objective counterpart, i.e., TS. As was explained in Section VI-A, we show the three representative points in Table VIII in the accuracy–semantic and the accuracy–complexity objective planes. This

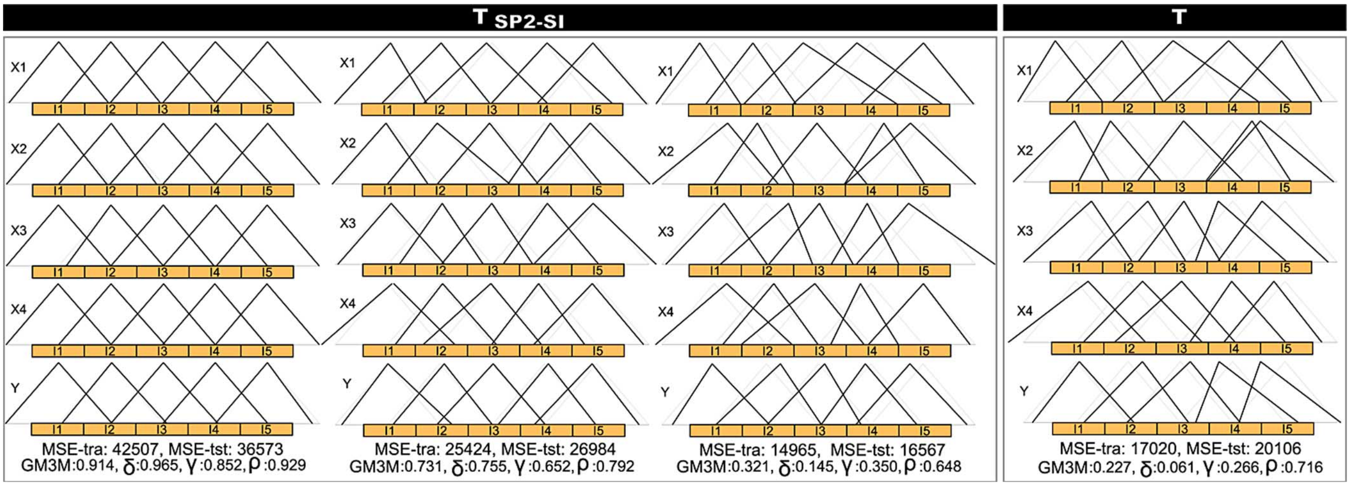


Fig. 3. DB obtained with T and three representative DBs obtained with $T_{SP2\text{-}SI}$ from one run in ELE.

TABLE VI
RESULTS OBTAINED BY THE METHODS THAT PERFORM ONLY RULE SELECTION

Dataset	Method	*MAX* INT			MEDIAN (INT/ACC)			*MAX* ACC		
		NR	$MSE^{tra}_{/2}$	$MSE^{tst}_{/2}$	NR	$MSE^{tra}_{/2}$	$MSE^{tst}_{/2}$	NR	$MSE^{tra}_{/2}$	$MSE^{tst}_{/2}$
PLA	S		-			-		12.6	2.416	**2.416**
	S_{SP2}	6.0	7.042	7.042	9.0	3.473	3.473	**12.0**	2.416	**2.416**
QUA	S		-			-		30.3	0.0220	0.0321
	S_{SP2}	17.4	0.0251	0.0260	24.1	0.0222	0.0234	**28.6**	0.0220	**0.0234**
ELE	S		-			-		40.9	41184	43049
	S_{SP2}	23.7	75486	81764	32.2	44980	46692	**39.0**	41608	**42987**
ABA	S		-			-		**17.7**	4.818	**4.810**
	S_{SP2}	9.9	11.124	11.093	16.4	5.240	5.227	22.2	5.071	5.052
STP	S		-			-		**37.0**	2.532	2.610
	S_{SP2}	30.4	3.741	3.786	45.9	1.521	1.605	62.7	1.446	**1.540**
WAN	S		-			-		47.9	6.418	**7.363**
	S_{SP2}	15.2	12.721	13.564	26.7	6.715	7.966	**38.7**	6.350	7.741
WIZ	S		-			-		44.0	3.036	6.909
	S_{SP2}	14.0	7.333	7.819	24.5	3.261	3.814	**33.7**	3.029	**3.499**
MOR	S		-			-		19.2	0.157	**0.165**
	S_{SP2}	4.2	1.485	1.529	10.1	0.320	0.330	**17.2**	0.252	0.255
TRE	S		-			-		19.7	0.251	**0.257**
	S_{SP2}	4.3	3.091	3.060	11.0	0.396	0.411	**18.5**	0.326	0.342

TABLE VII
WILCOXON'S TEST: S (R^+) VERSUS S_{SP2} (R^-) ON NR AND $MSE^{tst}_{/2}$ AT MAX ACC

Comparison	Measure	R^+	R^-	Hypothesis ($\alpha = 0.1$)	p-value
S vs. S_{SP2}	NR	17	28	Accepted	0.515
S vs. S_{SP2}	$MSE^{tst}_{/2}$	23	22	Accepted	0.953

allows further comparisons with the approaches that perform only rule selection and those that perform only tuning. In both cases, the values at the point MAX ACC coincide. The results of the single-objective counterpart algorithm, i.e., TS, are also shown in this table.

This time, we can compare the results from $TS_{SP2\text{-}SI}$ and TS on the three objective measures. Table IX shows the results of Wilcoxon's test for the most accurate point MAX ACC on them. For each measure, $TS_{SP2\text{-}SI}$ clearly outperforms TS. The null hypothesis for Wilcoxon's test in all the cases has been rejected in favor of $TS_{SP2\text{-}SI}$, with a very small *p*-value, which supports our conclusion with a high degree of confidence. It seems logical that by including NR and GM3M in the multiobjective approach, the interpretability should be better in the obtained FRBSs. However, they are also better in the accuracy objective. The use of the different measures to obtain a set of solutions with different tradeoffs helps maintain a higher diversity that promotes the derivation of more optimal solutions. Therefore, from these results and the results in the previous sections, we can conclude that in the approaches that consider tuning, it is preferable to use a multiobjective approach, including the proposed interpretability measures since we can obtain more interpretable and more accurate FRBSs than those obtained by the single-objective accuracy-oriented counterpart algorithms.

In Fig. 4, we represent some DBs that are obtained with TS and $TS_{SP2\text{-}SI}$ in ELE and PLA. See Section VI-B for an explanation of these kinds of figures. In both problems, it is clear that at least the DB with the best accuracy from $TS_{SP2\text{-}SI}$ is preferable to the one that is obtained by TS, but additional highly transparent DBs are also shown in the case of $TS_{SP2\text{-}SI}$.

E. Global Analysis on the Most Accurate Solutions: MAX ACC

Once the different approaches have been analyzed individually, all of them have to be compared to determine which of them should be preferred. In order to also include the single-objective-based algorithms, the global analysis is performed on the most accurate solutions. Since we will compare more than two algorithms, on this occasion, we use nonparametric tests for multiple comparisons. In order to perform a multiple comparison, it is necessary to check whether any of the results obtained by the algorithms present any inequality. In the case of finding some, we can know, by using a *post hoc* test, which algorithms partners' average results are dissimilar. We will use the results obtained in the evaluation of the three performance measures that have been presented in the previous sections, and we will define a control algorithm as the best performing algorithm (which obtains the lowest value of ranking that is computed through a Friedman test [58]). In order to test whether significant differences exist among all the mean values, we use Iman and Davenport's test [59]. Finally, we use Holm's [60] *post hoc* test to compare the control algorithm with what remains.

TABLE VIII
RESULTS OBTAINED BY THE METHODS THAT PERFORM BOTH RULE SELECTION AND TUNING OF MFS

Dataset	Method	Plane $MSE^{tst}_{/2}$,	*MAX* INT NR	$MSE^{tra}_{/2}$	$MSE^{tst}_{/2}$	GM3M	(δ	γ	ρ)	MEDIAN (INT/ACC) NR	$MSE^{tra}_{/2}$	$MSE^{tst}_{/2}$	GM3M	(δ	γ	ρ)	*MAX* ACC NR	$MSE^{tra}_{/2}$	$MSE^{tst}_{/2}$	GM3M	(δ	γ	ρ)
PLA	TS				-							-					**13.3**	1.255	1.301	0.251	(0.09	0.28	0.74)
	TS_{SP2-SI}	GM3M	12.8	2.624	2.631	0.966	(0.97	0.95	0.98)	13.2	1.600	1.647	0.832	(0.85	0.80	0.84)	13.7	1.170	**1.227**	**0.535**	(0.40	0.59	0.71)
	TS_{SP2-SI}	NR	7.2	1.981	2.022	0.566	(0.46	0.59	0.75)	10.3	1.335	1.412	0.430	(0.25	0.56	0.71)							
QUA	TS				-							-					33.5	0.0173	0.0428	0.182	(0.04	0.27	0.71)
	TS_{SP2-SI}	GM3M	50.5	0.0245	0.0252	0.943	(0.94	0.93	0.96)	34.9	0.0185	0.0192	0.715	(0.69	0.63	0.84)	**27.2**	0.0173	**0.0182**	**0.275**	(0.10	0.34	0.72)
	TS_{SP2-SI}	NR	19.1	0.0176	0.0184	0.372	(0.18	0.43	0.73)	22.4	0.0174	0.0182	0.304	(0.14	0.35	0.73)							
ELE	TS				-							-					41.3	13387	17784	0.335	(0.14	0.39	0.74)
	TS_{SP2-SI}	GM3M	60.2	50703	51959	0.941	(0.94	0.93	0.95)	40.6	28131	31090	0.798	(0.79	0.77	0.84)	**29.3**	11611	**14851**	**0.528**	(0.41	0.58	0.70)
	TS_{SP2-SI}	NR	17.4	34375	38453	0.623	(0.54	0.62	0.73)	21.7	17619	22099	0.544	(0.43	0.55	0.71)							
ABA	TS				-							-					28.4	2.473	2.582	0.021	(0.08	0.36	0.74)
	TS_{SP2-SI}	GM3M	48.1	6.530	6.576	0.958	(0.95	0.96	0.97)	35.8	3.381	3.453	0.826	(0.83	0.77	0.89)	**16.3**	2.386	**2.513**	**0.450**	(0.29	0.54	0.68)
	TS_{SP2-SI}	NR	7.7	3.317	3.423	0.564	(0.47	0.53	0.73)	11.0	2.554	2.670	0.474	(0.35	0.53	0.69)							
STP	TS				-							-					45.6	0.674	1.194	0.276	(0.14	0.25	0.69)
	TS_{SP2-SI}	GM3M	102.0	7.222	7.243	0.964	(0.96	0.95	0.98)	41.3	2.351	2.409	0.808	(0.83	0.75	0.85)	**32.9**	0.642	**0.775**	**0.364**	(0.25	0.40	0.69)
	TS_{SP2-SI}	NR	14.7	2.012	2.194	0.521	(0.41	0.49	0.77)	20.7	0.923	1.064	0.415	(0.28	0.41	0.72)							
WAN	TS				-							-					72.3	1.674	2.373	0.246	(0.08	0.33	0.70)
	TS_{SP2-SI}	GM3M	128.7	12.779	13.704	0.951	(0.94	0.94	0.97)	94.7	5.114	5.760	0.838	(0.86	0.80	0.86)	**39.3**	1.292	**2.016**	**0.456**	(0.31	0.52	0.72)
	TS_{SP2-SI}	NR	19.6	3.315	3.911	0.540	(0.41	0.56	0.74)	26.4	1.685	2.436	0.489	(0.36	0.49	0.72)							
WIZ	TS				-							-					53.5	1.051	2.386	0.349	(0.16	0.45	0.71)
	TS_{SP2-SI}	GM3M	96.5	6.216	6.448	0.962	(0.96	0.96	0.97)	67.7	2.630	2.680	0.833	(0.82	0.82	0.86)	**29.2**	0.921	**1.095**	**0.493**	(0.34	0.57	0.70)
	TS_{SP2-SI}	NR	9.9	2.599	2.963	0.543	(0.42	0.56	0.73)	15.9	1.361	1.522	0.530	(0.40	0.57	0.73)							
MOR	TS				-							-					34.1	0.031	0.037	0.316	(0.14	0.36	0.76)
	TS_{SP2-SI}	GM3M	55.2	0.608	0.634	0.937	(0.93	0.93	0.96)	29.3	0.168	0.180	0.851	(0.85	0.81	0.89)	**15.4**	0.028	**0.034**	**0.541**	(0.44	0.50	0.76)
	TS_{SP2-SI}	NR	5.1	0.202	0.213	0.587	(0.51	0.53	0.76)	8.8	0.065	0.073	0.538	(0.42	0.53	0.75)							
TRE	TS				-							-					29.9	0.050	0.065	0.319	(0.14	0.33	0.76)
	TS_{SP2-SI}	GM3M	57.5	1.141	1.155	0.957	(0.95	0.95	0.97)	32.1	0.310	0.324	0.866	(0.87	0.83	0.90)	**17.7**	0.040	**0.048**	**0.533**	(0.42	0.51	0.74)
	TS_{SP2-SI}	NR	4.9	0.436	0.448	0.583	(0.51	0.54	0.75)	8.9	0.081	0.083	0.531	(0.41	0.50	0.74)							

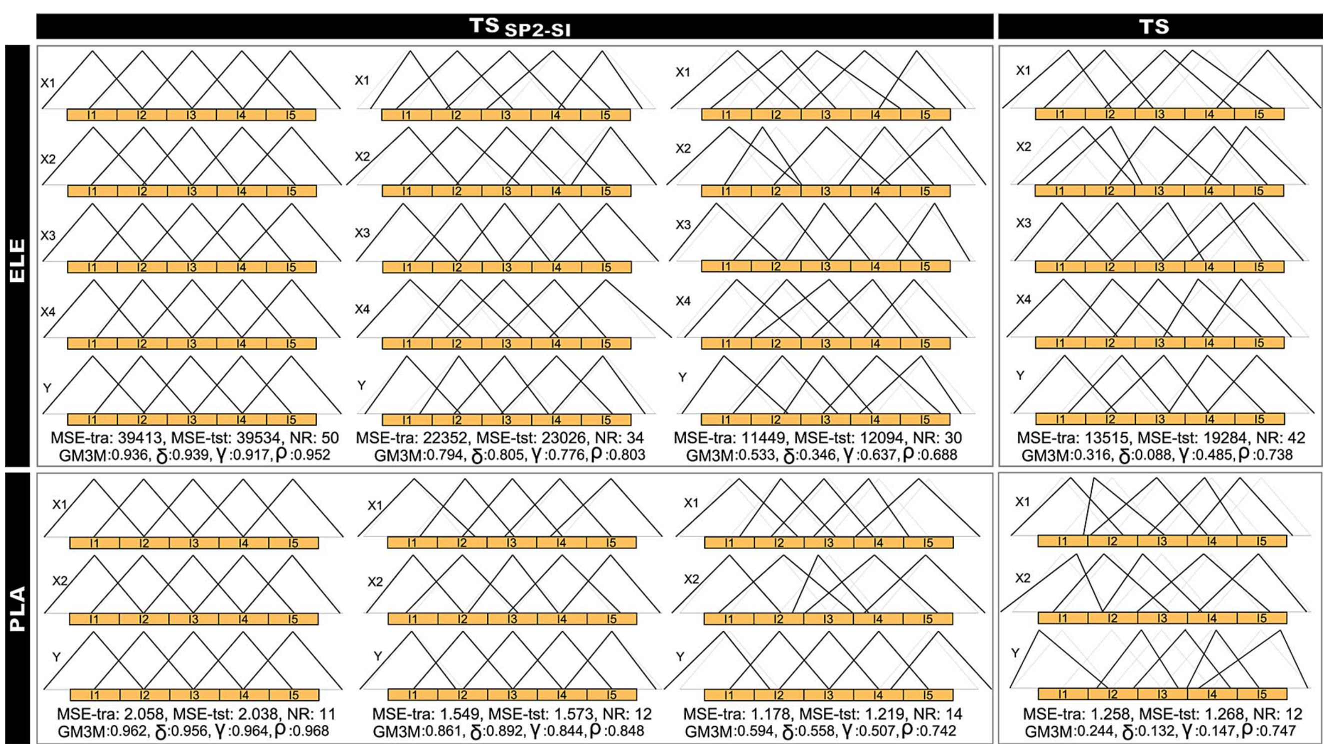


Fig. 4. DB obtained with TS and three representative DBs obtained with $TS_{SP2\text{-}SI}$ from one run in ELE and PLA.

TABLE IX
WILCOXON'S TEST: TS (R^+) VERSUS $TS_{SP2\text{-}SI}$ (R^-) ON GM3M, NR AND $MSE^{tst}_{/2}$ AT MAX ACC

Comparison	Measure	R^+	R^-	Hypothesis ($\alpha = 0.1$)	p-value
TS vs. TS_{SP2-SI}	GM3M	0	45	Rejected	0.008
TS vs. TS_{SP2-SI}	NR	1	44	Rejected	0.011
TS vs. TS_{SP2-SI}	$MSE^{tst}_{/2}$	0	45	Rejected	0.008

TABLE X
RANKINGS OBTAINED THROUGH FRIEDMAN'S TEST FOR THE METHODS THAT PERFORM SELECTION ON $MSE^{tst}_{/2}$ AND NR MEASURES

Algorithm	Ranking on $MSE^{tst}_{/2}$	Ranking on NR
TS_{SP2-SI}	1.0000	1.4444
TS	2.2222	3.7778
S_{SP2}	3.4444	2.1111
S	3.3333	2.6667

TABLE XI
HOLM TABLE FOR THE METHODS THAT PERFORM SELECTION WITH $\alpha = 0.1$ ON $\mathrm{MSE}^{\mathrm{tst}}_{/2}$ AND NR MEASURES

Holm's post-hoc test on $MSE^{tst}_{/2}$						Holm's post-hoc test on NR					
i	Algorithm	z	p	α/i	Hypothesis	i	Algorithm	z	p	α/i	Hypothesis
3	S_{SP2}	4.02	5.90E-5	0.03	Rejected	3	TS	3.83	1.26E-4	0.03	Rejected
2	S	3.83	1.26E-4	0.05	Rejected	2	S	2.01	0.045	0.05	Rejected
1	TS	2.01	0.05	0.1	Rejected	1	S_{SP2}	1.1	0.27	0.1	Accepted

TABLE XII
RANKINGS OBTAINED THROUGH FRIEDMAN'S TEST FOR THE METHODS THAT PERFORM TUNING ON $\mathrm{MSE}^{\mathrm{tst}}_{/2}$ AND GM3M MEASURES

Algorithm	Ranking on $MSE^{tst}_{/2}$	Ranking on GM3M
TS_{SP2-SI}	1.0000	1.0000
TS	3.2222	2.4444
T_{SP2-SI}	2.5555	3.0
T	3.2222	3.5555

As explained in Section VI-A, the approaches that consider rule selection should be compared in the accuracy–complexity plane, while the approaches that consider tuning should be compared in the accuracy–semantic plane. For this reason, we perform two studies: the first one on the methods that perform rule selection and the second one on the methods that perform tuning. Obviously, TS and the proposed approach, i.e., $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$, are included in both studies by using their projections (see Section VI-A).

1) Analysis of the Methods That Perform Rule Selection—Accuracy–Complexity Plane: Table X shows the rankings of the different methods that are considered in this study. Iman–Davenport's test tells us that significant differences exist among the results observed in all datasets, with p-values (3.990E$-$8) and (8.214E$-$5) on $\mathrm{MSE}^{\mathrm{tst}}_{/2}$ and NR, respectively. The best ranking is obtained by $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ in both measures: test error and NR.

We now apply Holm's test to compare the best ranking method in each case with the remaining methods. Table XI presents these results, where, the algorithms are ordered with respect to the obtained z-value. Holm's test rejects the hypothesis of equality with the rest of the methods in $\mathrm{MSE}^{\mathrm{tst}}_{/2}$ ($p < \alpha/i$). It also rejects the hypothesis with TS and S in NR. From these results, we can state that $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ outperforms the remaining methods in both accuracy and complexity, except in the case of $\mathrm{S}_{\mathrm{SP2}}$ that should be considered to be equivalent in terms of NR. However, we can ensure that under these conditions, $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ dominates $\mathrm{S}_{\mathrm{SP2}}$. It is also interesting to note the ranking position that is obtained by TS on NR. It shows that some unnecessary or inadequate rules cannot be removed by the single-objective approach.

2) Analysis of the Methods That Perform Tuning—Accuracy–Semantic Plane: In this study, Table XII shows the rankings (through Friedman's test) of the four algorithms considered. The p-values computed using Iman–Davenport's test [(8.171E$-$6) and (6.956E$-$7)] imply that there are statistical differences among the results on $\mathrm{MSE}^{\mathrm{tst}}_{/2}$ and GM3M, respectively. $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ is better in ranking for both measures. In both cases, Holm's test (see Table XIII) rejects the null hypothesis with all the remaining methods. The best method is again $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$, which obtains the best results for these two objectives. Finally, since the proposed approach is the best in both planes, we can conclude that this method is preferable to the remaining approaches to obtain accurate and simple FRBSs, thus maintaining a good level of semantic interpretability.

F. Graphical and Statistical Analysis of the Pareto Fronts

Since we perform 30 trials with different training and test partitions, it would not be readable to show all the Pareto fronts. Thus, to have a glimpse of the trends of the Pareto fronts in the accuracy–complexity and the accuracy–semantic planes, we plot the MAX INT, the MEDIAN (INT/ACC), and the MAX ACC points for each MOEA and for each dataset in Fig. 5. We also show the solutions that are generated by the single-objective methods.

The analysis of Fig. 5 shows that the approximations of the Pareto fronts that are achieved by $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ are, in general, below the approximations of the Pareto fronts that are obtained by the other MOEAs. To compare in detail the different MOEAs with respect to the MAX INT and MEDIAN (INT/ACC) points, we show the results of the application of the Wilcoxon test on these points in Table XIV for the MOEAs that perform rule selection, i.e., $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ and $\mathrm{S}_{\mathrm{SP2}}$. We observe a behavior that is very similar to the MAX ACC point, i.e., $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ outperforms $\mathrm{S}_{\mathrm{SP2}}$ in all the cases, except for NR in the most interpretable point.

With regard to the MOEAs that perform tuning, we show the results of the application of the Wilcoxon test for the same points in Table XV. At the MEDIAN (INT/ACC) point, the null hypothesis that is associated with the Wilcoxon test is rejected ($p < \alpha$) in GM3M, although the results achieved by $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ and $\mathrm{T}_{\mathrm{SP2\text{-}SI}}$ are statistically equivalent on $\mathrm{MSE}^{\mathrm{tst}}_{/2}$, which is the same as that obtained with MAX INT, but the role between both measures is changed. Under these conditions, we can state that the solutions that are obtained by $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ dominate, in general, the ones that are obtained by $\mathrm{T}_{\mathrm{SP2\text{-}SI}}$ in practically all the parts of the Pareto fronts.

In order to show the actual behavior of the approximated Pareto fronts provided by each MOEA, we show some representative Pareto fronts (the results of a single trial) on two datasets in Fig. 6. In this figure, we plot the solutions from $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ in a 3-D way, and we plot the projections of these solutions on all the possible objective planes along with the corresponding comparison methods. In order to retain all the information, the dominated solutions that are obtained from the projections have not been removed. The symbols and colors similar to those used in Fig. 5 have been used in this case.

TABLE XIII
HOLM TABLE FOR THE METHODS THAT PERFORM TUNING WITH $\alpha = 0.1$ ON $\text{MSE}^{\text{tst}}_{/2}$ AND GM3M MEASURES

Holm's post-hoc test on $MSE^{tst}_{/2}$						Holm's post-hoc test on GM3M					
i	Algorithm	z	p	α/i	Hypothesis	i	Algorithm	z	p	α/i	Hypothesis
3	T	3.65	2.61E-4	0.03	Rejected	3	T	4.20	2.68E-5	0.03	Rejected
2	TS	3.65	2.61E-4	0.05	Rejected	2	T_{SP2-SI}	3.29	0.001	0.05	Rejected
1	T_{SP2-SI}	2.56	0.01	0.1	Rejected	1	TS	2.37	0.018	0.1	Rejected

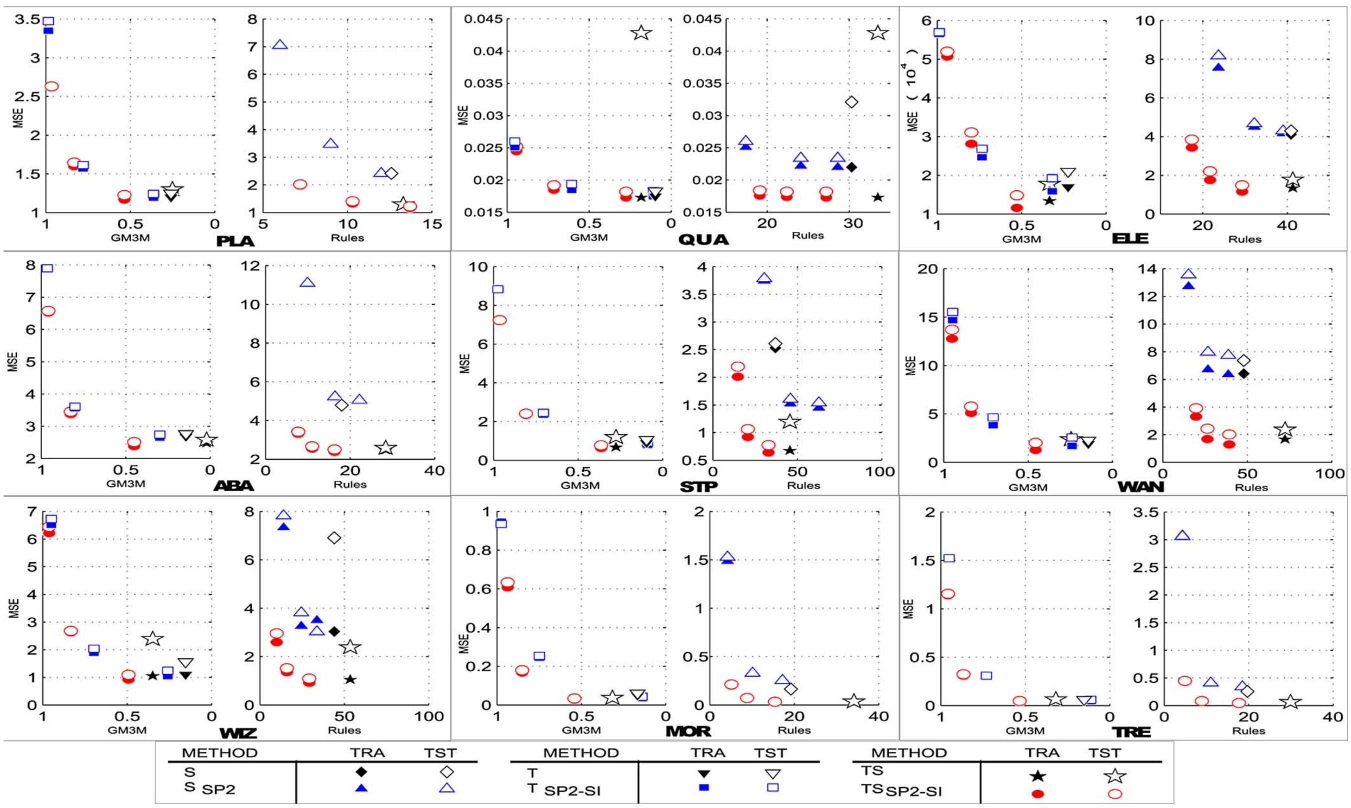


Fig. 5. Averaged Pareto fronts that are obtained in all the problems.

TABLE XIV
WILCOXON'S TEST: S_{SP2} (R^+) VERSUS $\text{TS}_{\text{SP2-SI}}$ (R^-) ON NR AND $\text{MSE}^{\text{tst}}_{/2}$ AT MEDIAN (INT/ACC) AND MAX INT

Comparison	Measure	Pareto solution	R^+	R^-	Hypothesis ($\alpha = 0.1$)	p-value
S_{SP2} vs. TS_{SP2-SI}	NR	MEDIAN (INT/ACC)	3	42	Rejected	0.021
S_{SP2} vs. TS_{SP2-SI}	$MSE^{tst}_{/2}$	MEDIAN (INT/ACC)	1	44	Rejected	0.011
S_{SP2} vs. TS_{SP2-SI}	NR	MAX INT	25	20	Accepted	0.767
S_{SP2} vs. TS_{SP2-SI}	$MSE^{tst}_{/2}$	MAX INT	0	45	Rejected	0.008

TABLE XV
WILCOXON'S TEST: $\text{T}_{\text{SP2-SI}}$ (R^+) VERSUS $\text{TS}_{\text{SP2-SI}}$ (R^-), GM3M AND $\text{MSE}^{\text{tst}}_{/2}$ AT MEDIAN (INT/ACC) AND MAX INT

Comparison	Measure	Pareto solution	R^+	R^-	Hypothesis ($\alpha = 0.1$)	p-value
T_{SP2-SI} vs. TS_{SP2-SI}	GM3M	MEDIAN (INT/ACC)	0	45	Rejected	0.008
T_{SP2-SI} vs. TS_{SP2-SI}	$MSE^{tst}_{/2}$	MEDIAN (INT/ACC)	30	15	Accepted	0.374
T_{SP2-SI} vs. TS_{SP2-SI}	GM3M	MAX INT	36	9	Accepted	0.110
T_{SP2-SI} vs. TS_{SP2-SI}	$MSE^{tst}_{/2}$	MAX INT	0	45	Rejected	0.008

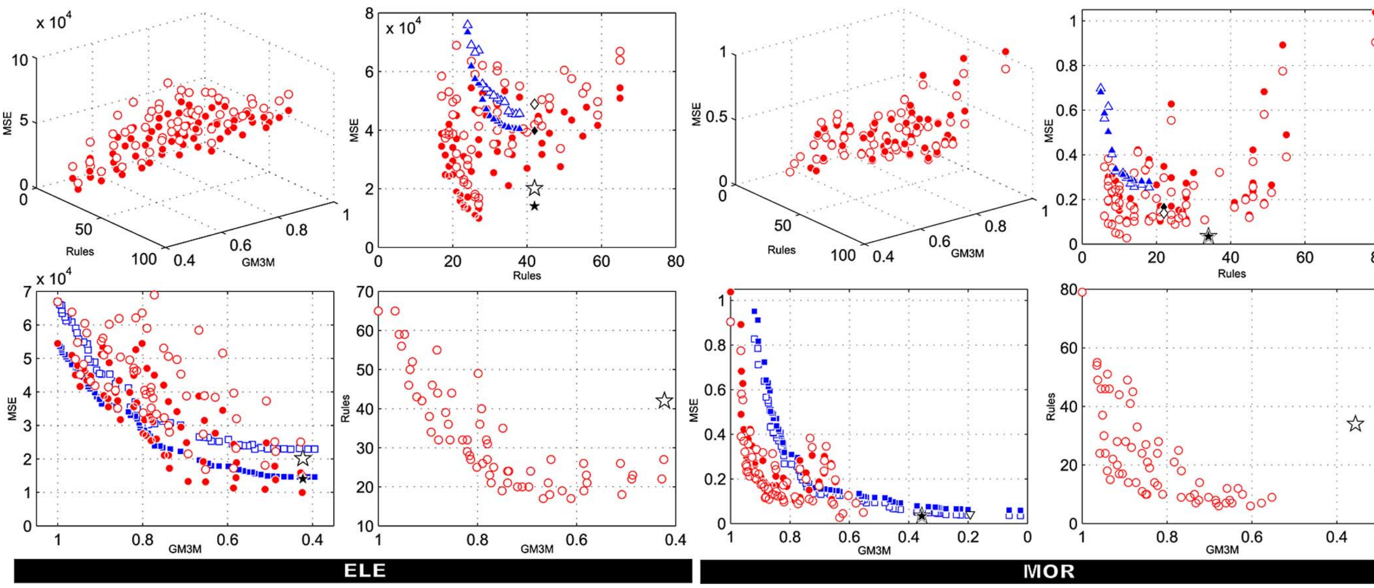


Fig. 6. Example Pareto fronts that are obtained in ELE and MOR problems.

VII. Conclusion

In this paper, we have proposed an index that helps preserve the semantic interpretability of linguistic fuzzy systems, namely, GM3M. The GM3M index is devoted to preserving the original shape of the MFs while a tuning of their definition parameters is performed, and it represents a measure of the quality of the DB. It works on the assumption that the initial DB comprises the appropriate MFs with an associated linguistic meaning (which is usually given by an expert). To this end, we have proposed $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$, which is an effective postprocessing MOEA that is designed to generate a set of FRBSs with different tradeoffs among accuracy, complexity, and semantic interpretability. Three criteria have been considered: the $\mathrm{MSE}_{/2}$, the NR, and the proposed GM3M index. This method performs rule selection and tuning of the MFs simultaneously on a given initial linguistic FRBS.

We have shown that the use of the GM3M index within a multiobjective evolutionary framework helps the tuning approaches obtain more interpretable and, at the same time, more accurate models. Therefore, a multiobjective framework allows us to obtain FRBSs that are characterized by better tradeoffs between accuracy, complexity, and semantic interpretability than the ones that are provided by considering only accuracy as the unique objective.

We should point out that the interaction of rule selection with the tuning of MFs enables the derivation of much more accurate models, while at the same time, the semantic interpretability is maintained to a higher extent. Rule selection allows a major reduction in the system's complexity. Further, we observe that $\mathrm{TS}_{\mathrm{SP2\text{-}SI}}$ outperforms all the analyzed methods in all the datasets on the test error, and it achieves better values in GM3M when performing a tuning of the MFs. This way, very interesting solutions have also been obtained with improved accuracy and very high levels of semantic interpretability (near the initial model).

In this sense, this paper has proposed an index to measure the interpretability that is associated with the fuzzy partition along with an RB postprocessing method for obtaining a tradeoff between accuracy and interpretability in linguistic modeling. Working this way follows the final goal pursued by CW by improving the granulation of a continuous variable, which involves a partitioning of the whole into parts, while keeping the meaning of the original words and decreasing the complexity of the RB.

Appendix

On the Use of Nonparametric Tests Based on Rankings

A nonparametric test uses either nominal data, ordinal data, or data represented in an ordinal way of ranking. This does not imply that only they can be used for these types of data. It could be very interesting to transform the data from real values that are contained within an interval to ranking-based data, which is similar to the way a nonparametric test can be applied over typical data of a parametric test when they do not fulfill the necessary conditions that are imposed by the use of the test. In the following, we explain the basic functionality of each nonparametric test used in this study, along with the aim that is pursued by its use.

1) Friedman's test [58]: It is a nonparametric equivalent of the test of repeated-measures analysis of variance (ANOVA). It computes the ranking of the observed results for algorithm (r_j for the algorithm j with k algorithms) for each dataset, assigning the ranking 1 to the best of them and the ranking k to the worst. Under the null hypothesis, which is

formed by assuming that the results of the algorithms are equivalent (with similar rankings), Friedman's statistic

$$\mathcal{X}_F^2 = \frac{12N_{ds}}{k(k+1)}\left[\sum_j R_j^2 - \frac{k(k+1)^2}{4}\right] \tag{10}$$

is distributed according to $\mathcal{X}_F^2$ with $k-1$ degrees of freedom (DOFs), where $R_j = (1/N_{ds})\sum_i R_i^j$, and N_{ds} is the number of datasets. The critical values for Friedman's statistic coincide with those established in the $\mathcal{X}^2$-distribution when $N_{ds} > 10$ and $k > 5$. On the contrary, the exact values can be seen in [56] and [61].

2) Iman and Davenport's test [59]: It is a metric that is derived from Friedman's statistic given that this last metric produces a conservative undesirable effect. The statistic is

$$\mathcal{F}_F = \frac{(N_{ds}-1)\mathcal{X}_F^2}{N_{ds}(k-1) - \mathcal{X}_F^2} \tag{11}$$

and it is distributed as an *F*-distribution with $k-1$ and $(k-1)(N_{ds}-1)$ DOFs.

3) Holm's method [60]: This test sequentially checks the hypothesis ordered according to their significance. We will denote the *p*-values ordered by $p_1, p_2, \ldots$ in such a way that $p_1 \le p_2 \le \cdots \le p_{k-1}$. Holm's method compares each p_i with $\alpha/(k-i)$, starting from the most significant *p*-value. If p_1 is less than $\alpha/(k-1)$, the corresponding hypothesis is rejected, and it allows the comparison of p_2 with $\alpha/(k-2)$. If the second hypothesis is rejected, we continue with the process. As soon as a certain hypothesis cannot be rejected, all the remaining hypotheses are maintained as accepted. The statistic for comparing the i algorithm with the j algorithm is

$$z = \frac{(R_i - R_j)}{\sqrt{(k(k+1))/6N_{ds}}}. \tag{12}$$

The value of z is used to find the corresponding probability from the table of the normal distribution, which is compared with the corresponding value of α.

4) Wilcoxon's signed-rank test: The Wilcoxon signed-rank test is a pairwise test with the aim of detecting significant differences between two sample means: It is analogous to the paired *t*-test in nonparametric statistical procedures. If these means refer to the outputs of two algorithms, then the test practically assesses the reciprocal behavior of the two algorithms [56], [57]. Let d_i be the difference between the performance scores of the two algorithms on the *i*th out of N_{ds} datasets. The differences are ranked according to their absolute values; average ranks are assigned in case of ties. Let R^+ be the sum of ranks for the datasets on which the first algorithm outperformed the second, and let R^- be the sum of ranks for the contrary outcome. Ranks of $d_i = 0$ are split evenly among the sums; if there is an odd number of them, one is ignored:

$$R^+ = \sum_{d_i>0} \text{rank}(d_i) + \frac{1}{2}\sum_{d_i=0}\text{rank}(d_i)$$
$$R^- = \sum_{d_i<0} \text{rank}(d_i) + \frac{1}{2}\sum_{d_i=0}\text{rank}(d_i). \tag{13}$$

Let T be the smaller of the sums, i.e., $T = \min(R^+, R^-)$. If T is either less than or equal to the value of the distribution of Wilcoxon for N_{ds} DOFs (see [61, Tab. B.12]), the null hypothesis of equality of means is rejected.

REFERENCES

[1] L. A. Zadeh, "From computing with numbers to computing with words—From manipulation of measurements to manipulation of perceptions," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 45, no. 1, pp. 105–119, Jan. 1999.

[2] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Int. J. Man–Mach. Stud.*, vol. 7, pp. 1–13, 1975.

[3] E. H. Mamdani, "Application of fuzzy logic to approximate reasoning using linguistic systems," *IEEE Trans. Comput.*, vol. C-26, no. 12, pp. 1182–1191, Dec. 1977.

[4] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning, Part I," *Inf. Sci.*, vol. 8, pp. 199–249, 1975.

[5] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning, Part II," *Inf. Sci.*, vol. 8, pp. 301–357, 1975.

[6] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning, Part III," *Inf. Sci.*, vol. 9, pp. 43–80, 1975.

[7] J. M. Alonso, L. Magdalena, and G. Gonzalez-Rodriguez, "Looking for a good fuzzy system interpretability index. An experimental approach," *Int. J. Approx. Reason.*, vol. 51, no. 1, pp. 115–134, 2009.

[8] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena, Eds., *Accuracy Improvements in Linguistic Fuzzy Modeling* (Studies in Fuzziness and Soft Computing, vol. 129). New York: Springer-Verlag, 2003.

[9] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena, Eds., *Interpretability Issues in Fuzzy Modeling* (Studies in Fuzziness and Soft Computing, vol. 128). New York: Springer-Verlag, 2003.

[10] C. Mencar and A. Fanelli, "Interpretability constraints for fuzzy information granulation," *Inf. Sci.*, vol. 178, no. 24, pp. 4585–4618, 2008.

[11] S. M. Zhou and J. Q. Gan, "Low-level interpretability and high-level interpretability: A unified view of data-driven interpretable fuzzy system modelling," *Fuzzy Sets Syst.*, vol. 159, no. 23, pp. 3091–3131, 2008.

[12] J. M. Alonso, L. Magdalena, and S. Guillaume, "Hilk: A new methodology for designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism," *Int. J. Intell. Syst.*, vol. 23, no. 7, pp. 761–794, 2008.

[13] R. Alcalá, M. J. Gacto, F. Herrera, and J. Alcalá-Fdez, "A multi-objective genetic algorithm for tuning and rule selection to obtain accurate and compact linguistic fuzzy rule-based systems," *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 15, no. 5, pp. 539–557, 2007.

[14] M. Cococcioni, P. Ducange, B. Lazzerini, and F. Marcelloni, "A Pareto-based multi-objective evolutionary approach to the identification of Mamdani fuzzy systems," *Soft Comput.*, vol. 11, pp. 1013–1031, 2007.

[15] M. J. Gacto, R. Alcalá, and F. Herrera, "Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems," *Soft Comput.*, vol. 13, no. 5, pp. 419–436, 2009.

[16] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, "Selecting fuzzy if–then rules for classification problems using genetic algorithms," *IEEE Trans. Fuzzy Syst.*, vol. 3, no. 3, pp. 260–270, Aug. 1995.

[17] H. Ishibuchi, T. Murata, and I. B. Türksen, "Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems," *Fuzzy Sets Syst.*, vol. 89, no. 2, pp. 135–150, 1997.

[18] H. Ishibuchi, T. Nakashima, and T. Murata, "Three-objective genetics-based machine learning for linguistic rule extraction," *Inf. Sci.*, vol. 136, pp. 109–133, 2001.

[19] H. Ishibuchi and T. Yamamoto, "Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining," *Fuzzy Sets Syst.*, vol. 141, no. 1, pp. 59–88, 2004.
[20] H. Ishibuchi and Y. Nojima, "Analysis of interpretability–accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning," *Int. J. Approx. Reason.*, vol. 44, no. 1, pp. 4–31, 2007.
[21] P. Pulkkinen and H. Koivisto, "Fuzzy classifier identification using decision tree and multiobjective evolutionary algorithms," *Int. J. Approx. Reason.*, vol. 48, no. 2, pp. 526–543, 2008.
[22] U. Bodenhofer and P. Bauer, "A formal model of interpretability of linguistic variables," in *Interpretability Issues in Fuzzy Modeling*, J. Casillas, O. Cordón, F. Herrera, and L. Magdalena, Eds. New York: Springer-Verlag, 2003, pp. 524–545.
[23] A. Botta, B. Lazzerini, F. Marcelloni, and D. Stefanescu, "Context adaptation of fuzzy systems through a multiobjective evolutionary approach based on a novel interpretability index," *Soft Comput.*, vol. 13, no. 5, pp. 437–449, 2009.
[24] F. Cheong and R. Lai, "Constraining the optimization of a fuzzy logic controller using an enhanced genetic algorithm," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 30, no. 1, pp. 31–46, Feb. 2000.
[25] J. Espinosa and J. Vandewalle, "Constructing fuzzy models with linguistic integrity from numerical data—AFRELI algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 5, pp. 591–600, Oct. 2000.
[26] P. Fazendeiro, J. V. de Oliveira, and W. Pedrycz, "A multiobjective design of a patient and anaesthetist-friendly neuromuscular blockade controller," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 9, pp. 1667–1678, Sep. 2007.
[27] Y. Jin, W. von Seelen, and B. Sendhoff, "On generating fc^3 fuzzy rule systems from data using evolution strategies," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 29, no. 6, pp. 829–845, Dec. 1999.
[28] C. Mencar, G. Castellano, and A. M. Fanelli, "Distinguishability quantification of fuzzy sets," *Inf. Sci.*, vol. 177, pp. 130–149, 2007.
[29] D. Nauck, "Measuring interpretability in rule-based classification systems," in *Proc. 12th IEEE Int. Conf. Fuzzy Syst.,* 2003, vol. 1, no. 2, pp. 196–201.
[30] M. Setnes, R. Babuška, U. Kaymak, and H. R. van Nauta Lemke, "Similarity measures in fuzzy rule base simplification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 28, no. 3, pp. 376–386, Jun. 1998.
[31] J. V. de Oliveira, "Semantic constraints for membership function optimization," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 29, no. 1, pp. 128–138, Jan. 1999.
[32] J. V. de Oliveira, "Towards neuro-linguistic modeling: Constraints for optimization of membership functions," *Fuzzy Sets Syst.*, vol. 106, no. 3, pp. 357–380, 1999.
[33] R. Alcalá, J. Alcalá-Fdez, and F. Herrera, "A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection," *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 4, pp. 616–635, Aug. 2007.
[34] R. Alcalá, J. Alcalá-Fdez, M. J. Gacto, and F. Herrera, "Rule base reduction and genetic tuning of fuzzy systems based on the linguistic 3-tuples representation," *Soft Comput.*, vol. 11, no. 5, pp. 401–419, 2007.
[35] J. Casillas, O. Cordón, M. J. del Jesus, and F. Herrera, "Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction," *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 1, pp. 13–29, Feb. 2005.
[36] O. Cordón and F. Herrera, "A three-stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples," *Int. J. Approx. Reason.*, vol. 17, no. 4, pp. 369–407, 1997.
[37] H. B. Gürocak, "A genetic-algorithm-based method for tuning fuzzy logic controllers," *Fuzzy Sets Syst.*, vol. 108, no. 1, pp. 39–47, 1999.
[38] F. Herrera, M. Lozano, and J. L. Verdegay, "Tuning fuzzy logic controllers by genetic algorithms," *Int. J. Approx. Reason.*, vol. 12, pp. 299–315, 1995.
[39] C. L. Karr, "Genetic algorithms for fuzzy controllers," *AI Expert*, vol. 6, no. 2, pp. 26–33, 1991.
[40] O. Cordón, F. Herrera, F. Hoffmann, and L. Magdalena, *Genetic Fuzzy Systems. Evolutionary Tuning and Learning of Fuzzy Knowledge Base* (Advances in Fuzzy Systems—Applications and Theory, vol. 19). Singapore: World Scientific, 2001.
[41] C. A. Coello, D. A. V. Veldhuizen, and G. B. Lamont, Eds., *Evolutionary Algorithms for Solving Multi-Objective Problems*. Norwell, MA: Kluwer, 2002.
[42] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. New York: Wiley, 2001.
[43] F. Herrera, M. Lozano, and J. L. Verdegay, "A learning process for fuzzy control rules using genetic algorithms," *Fuzzy Sets Syst.*, vol. 100, pp. 143–158, 1998.
[44] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization," in *Proc. Evol. Methods Des., Optim. Control Appl. Ind. Probl.*, Barcelona, Spain, 2001, pp. 95–100.
[45] L. J. Eshelman, "The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination," in *Foundations of GAs*, vol. 1, G. Rawlin, Ed., San Mateo, CA: Morgan Kaufman, 1991, pp. 265–283.
[46] K. Deb, A. Pratab, S. Agrawal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
[47] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *J. Mach. Learn. Res.*, vol. 7, pp. 1–30, 2006.
[48] S. García and F. Herrera, "An extension on statistical comparisons of classifiers over multiple data sets for all pairwise comparisons," *J. Mach. Learn. Res.*, vol. 9, pp. 2677–2694, 2008.
[49] S. García, A. Fernández, J. Luengo, and F. Herrera, "A study of statistical techniques and performance measures for genetics-based machine learning: Accuracy and interpretability," *Soft Comput.*, vol. 13, no. 10, pp. 959–977, 2009.
[50] S. García, D. Molina, M. Lozano, and F. Herrera, "A study on the use of non-parametric tests for analyzing the evolutionary algorithms' behaviour: A case study on the CEC'2005 special session on real parameter optimization," *J. Heuristics*, vol. 15, pp. 617–644, 2009.
[51] A. Krone and H. Taeger, "Data-based fuzzy rule test for fuzzy modelling," *Fuzzy Sets Syst.*, vol. 123, no. 3, pp. 343–358, 2001.
[52] E. H. Ruspini, "A new approach to clustering," *Inf. Control*, vol. 15, no. 1, pp. 22–32, 1969.
[53] L. J. Eshelman and J. D. Schaffer, "Real-coded genetic algorithms and interval-schemata," *Found. Genet. Algorithms*, vol. 2, pp. 187–202, 1993.
[54] L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, no. 6, pp. 1414–1427, Nov./Dec. 1992.
[55] J. Alcalá-Fdez, L. Sánchez, S. García, M. del Jesus, S. Ventura, J. Garrell, J. Otero, C. Romero, J. Bacardit, V. Rivas, J. Fernández, and F. Herrera, "KEEL: A software tool to assess evolutionary algorithms to data mining problems," *Soft Comput.*, vol. 13, no. 3, pp. 307–318, 2009.
[56] D. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*. London, U.K./Boca Raton, FL: Chapman & Hall/CRC, 2003.
[57] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics*, vol. 1, pp. 80–83, 1945.
[58] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *J. Amer. Stat. Assoc.*, vol. 32, pp. 675–701, 1937.
[59] R. L. Iman and J. H. Davenport, "Approximations of the critical region of the Friedman statistic," *Commun. Stat. A, Theory Methods*, vol. 9, pp. 571–595, 1980.
[60] S. Holm, "A simple sequentially rejective multiple test procedure," *Scand. J. Stat.*, vol. 6, pp. 65–70, 1979.
[61] J. Zar, *Biostatistical Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1999.

María José Gacto received the M.Sc. degree in computer science from the University of Granada, Granada, Spain, in 1999.

She is currently with the Intelligent Systems and Data Mining Research Group, Department of Computer Science, University of Jaén, Jaén, Spain. She has authored or coauthored more than 20 international publications. She has worked on several research projects supported by the Spanish government and the European Union. Her current research interests include multiobjective genetic algorithms and genetic fuzzy systems, particularly the learning/tuning of fuzzy systems for modeling and control with a good tradeoff between accuracy and interpretability.

Rafael Alcalá received the M.Sc. and Ph.D. degrees in computer science from the University of Granada, Granada, Spain, in 1998 and 2003, respectively.

From 1998 to 2003, he was with the Department of Computer Science, University of Jaén, Jaén, Spain. He is currently an Assistant Professor with the Department of Computer Science and Artificial Intelligence, University of Granada, where he is also a member of the Soft Computing and Intelligent Information Systems Research Group. He has authored or coauthored more than 60 papers published in international journals, book chapters, and presented at conferences. His research interests include multiobjective genetic algorithms and genetic fuzzy systems, particularly the learning/tuning of fuzzy systems for modeling and control with a good tradeoff between accuracy and interpretability, as well as fuzzy association rules. He has coedited the special issue of *Evolutionary Intelligence* on "Genetic fuzzy systems: New advances." He has worked on several research projects supported by the Spanish government and the European Union. He is currently a member of the editorial/reviewer boards of several journals, including the *International Journal of Computational Intelligence Research*, the *Journal of Advanced Research in Fuzzy and Uncertain Systems*, and the *Journal of Universal Computer Science and Applied Intelligence*.

Dr. Alcalá has coedited the special issue of the IEEE TRANSACTIONS ON FUZZY SYSTEMS on "Genetic fuzzy systems: What's next."

Francisco Herrera received the M.Sc. and Ph.D. degrees in mathematics from the University of Granada, Granada, Spain, in 1988 and 1991, respectively.

He is currently a Professor with the Department of Computer Science and Artificial Intelligence, University of Granada. He has authored or coauthored more than 150 papers published in international journals. He is a coauthor of the book *Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Knowledge Bases* (World Scientific, 2001). He has coedited five international books and 20 special issues of international journals on different topics of soft computing. He is an Associate Editor of *Information Sciences*, *Mathware and Soft Computing*, *Advances in Fuzzy Systems*, *Advances in Computational Sciences and Technology*, and the *International Journal of Applied Metaheuristics Computing*. He is also an Area Editor of *Soft Computing* (area of genetic algorithms and genetic fuzzy systems) and is a member of the editorial boards of several journals, including *Fuzzy Sets and Systems*, *Applied Intelligence*, *Knowledge and Information Systems*, *Information Fusion*, *Evolutionary Intelligence*, the *International Journal of Hybrid Intelligent Systems*, and *Memetic Computation*. His current research interests include computing with words and decision making, data mining, data preparation, instance selection, fuzzy-rule-based systems, genetic fuzzy systems, knowledge extraction based on evolutionary algorithms, memetic algorithms, and genetic algorithms.

Prof. Herrera is an Associate Editor of the IEEE TRANSACTIONS ON FUZZY SYSTEMS.

Bibliografía

[AAFGH07] Alcalá R., Alcalá-Fdez J., Gacto M., y Herrera F. (2007) Rule base reduction and genetic tuning of fuzzy systems based on the linguistic 3-tuples representation. *Soft Computing* (11:5): 401–419.

[AAFH07] Alcalá R., Alcalá-Fdez J., y Herrera F. (2007) A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection. *IEEE Trans. Fuzzy Syst.* 15(4): 616–635.

[AFHMP07] Alcalá-Fdez J., Herrera F., Márquez F. A., y Peregrín A. (2007) Increasing fuzzy rules cooperation based on evolutionary adaptive inference systems. *International Journal of Intelligent Systems* 22(9): 1035–1064.

[AGHAF07] Alcalá R., Gacto M. J., Herrera F., y Alcalá-Fdez J. (2007) A multi-objective genetic algorithm for tuning and rule selection to obtain accurate and compact linguistic fuzzy rule-based systems. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 15(5): 539–557.

[AMG08] Alonso J. M., Magdalena L., y Guillaume S. (2008) Hilk: A new methodology for designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism. *International Journal of Intelligent Systems* 23(7): 761–794.

[AMR09] Alonso J. M., Magdalena L., y Rodríguez G. G. (2009) Looking for a good fuzzy system interpretability index: An experimental approach. *International Journal of Approximate Reasoning* 51: 115–134.

[BB03] Bodenhofer U. y Bauer P. (2003) A formal model of interpretability of linguistic variables. En Casillas J., Cordón O., Herrera F., y Magdalena L. (Eds.) *Interpretability issues in fuzzy modeling*, páginas 524–545. Springer-Verlag.

[BDDO04] Branke J., Deb K., Dierolf H., y Osswald M. (2004) Finding knees in multi-objective optimization. En *Proc. Parallel Problem Solving from Nature Conf. - PPSN VIII, LNCS 3242*, páginas 722–731. Birmingham, UK.

[BLMS09] Botta A., Lazzerini B., Marcelloni F., y Stefanescu D. C. (2009) Context adaptation of fuzzy systems through a multi-objective evolutionary approach based on a novel interpretability index. *Soft Computing* 13(5): 437–449.

[CBFO06] Crockett K., Bandar Z., Fowdar J., y O'Shea J. (2006) Genetic tuning of fuzzy inference within fuzzy classifier systems. *Expert Systems* 23(2): 63–82.

[CBM07] Crockett K., Bandar Z., y Mclean D. (2007) On the optimization of t-norm parameters within fuzzy decision trees. En *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE07).*

[CDLM07] Cococcioni M., Ducange P., Lazzerini B., y Marcelloni F. (2007) A pareto-based multi-objective evolutionary approach to the identification of mamdani fuzzy systems. *Soft Computing* 11: 1013–1031.

[CHV01] Cordón O., Herrera F., y Villar P. (2001) Generating the knowledge base of a fuzzy rule-based system by the genetic learning of data base. *IEEE Transactions on Fuzzy Systems* 9(4): 667–674.

[CJKO01] Corne D., Jerram N., Knowles J., y Oates M. (2001) Pesa-ii: Region based selection in evolutionary multiobjective optimization. En *Proc. Genetic and Evolutionary Computation Conf.*, páginas 283–290. San Francisco, CA.

[CKO00] Corne D., Knowles J., y Oates M. (2000) The pareto envelope-based selection algorithm for multiobjective optimization. En *Proc. Parallel Problem Solving from Nature Conf. - PPSN VI, LNCS 1917*, páginas 839–848. Paris, France.

[CL00] Cheong F. y Lai R. (2000) Constraining the optimization of a fuzzy logic controller using an enhanced genetic algorithm. *IEEE Trans. Syst., Man, Cybern. - Part B: Cybernetics* 30(1): 31–46.

[CT01] Coello C. y Toscano G. (2001) A micro-genetic algorithm for multiobjective optimization. En *First Int. Conf. on Evolutionary Multi-Criterion Optimization, LNCS 1993*, páginas 126–140. London, UK.

[dO99a] de Oliveira J. V. (1999) Semantic constraints for membership function optimization. *IEEE Trans. Syst., Man, Cybern. - Part A: Systems and Humans* 29(1): 128–138.

[dO99b] de Oliveira J. V. (1999) Towards neuro-linguistic modeling: constraints for optimization of membership functions. *Fuzzy Sets and Systems* 106(3): 357–380.

[DPAM02] Deb K., Pratab A., Agrawal S., y Meyarivan T. (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evolutionary Computation* 6(2): 182–197.

[EMH01] Erickson M., Mayer A., y Horn J. (2001) The niched pareto genetic algorithm 2 applied to the design of groundwater remediation systems. En *Proc. First Int. Conf. on Evolutionary Multi-Criterion Optimization, LNCS 1993*, páginas 681–695. London, UK.

[EV00] Espinosa J. y Vandewalle J. (2000) Constructing fuzzy models with linguistic integrity from numerical data-AFRELI algorithm. *IEEE Trans. Fuzzy Syst.* 8(5): 591–600.

[FdOP07] Fazendeiro P., de Oliveira J. V., y Pedrycz W. (2007) A multiobjective design of a patient and anaesthetist-friendly neuromuscular blockade controller. *IEEE Trans. on Biomedical Engineering* 54(9): 1667–1678.

[FF93] Fonseca C. y Fleming P. (1993) Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. En *Proc. 5th Int. Conf. on Genetic Algorithms*, páginas 416–423. San Mateo, CA.

[GAH09] Gacto M. J., Alcalá R., y Herrera F. (2009) Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems. *Soft Computing* 13(5): 419–436.

[GAH10] Gacto M., Alcalá R., y Herrera F. (2010) Integration of an index to preserve the semantic interpretability in the multi-objective evolutionary rule selection and tuning of linguistic fuzzy systems. *IEEE Transactions on Fuzzy Systems* 18(3): 515–531.

[Her08] Herrera F. (2008) Genetic fuzzy systems: taxonomy, current research trends and prospects. *Evolutionary Intelligence* 1: 27–46.

[HM95] Homaifar A. y McCormick E. (1995) Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms. *IEEE Transactions on Fuzzy Systems* 3(2): 129–139.

[HNG94] Horn J., Nafpliotis N., y Goldberg D. (1994) A niched pareto genetic algorithm for multiobjective optimization. En *Proc. First IEEE Conf. on Evolutionary Computation, IEEE World Congress on Computational Intelligence*, páginas 82–87. Piscataway, NJ.

[IMT97] Ishibuchi H., Murata T., y Türksen I. B. (1997) Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems. *Fuzzy Sets and Systems* 89(2): 135–150.

[IN07] Ishibuchi H. y Nojima Y. (2007) Analysis of interpretability-accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning. *International Journal of Approximate Reasoning* 44(1): 4–31.

[INM01] Ishibuchi H., Nakashima T., y Murata T. (2001) Three-objective genetics-based machine learning for linguistic rule extraction. *Information Sciences* 136: 109–133.

[INYT95] Ishibuchi H., Nozaki K., Yamamoto N., y Tanaka H. (1995) Selecting fuzzy if-then rules for classification problems using genetic algorithms. *IEEE Trans. Fuzzy Syst.* 3(3): 260–270.

[IY04] Ishibuchi H. y Yamamoto T. (2004) Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining. *Fuzzy Sets and Systems* 141(1): 59–88.

[JvSS99] Jin Y., von Seelen W., y Sendhoff B. (1999) On generating fc^3 fuzzy rule systems from data using evolution strategies. *IEEE Transactions on Systems, Man, and Cybernetics* 29(6): 829–845.

[KC00] Knowles J. D. y Corne D. W. (2000) Approximating the non dominated front using the pareto archived evolution strategy. *Evolutionary Computation* 8(2): 149–172.

[MA75] Mamdani E. H. y Assilian S. (1975) An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies* 7: 1–13.

[Mam74] Mamdani E. H. (1974) Application of fuzzy algorithms for control of simple dynamic plant. En *Proc. IEEE*, volumen 121, páginas 1585–1588.

[MCF07] Mencar C., Castellano G., y Fanelli A. M. (2007) Distinguishability quantification of fuzzy sets. *Information Sciences* 177: 130–149.

[MF08] Mencar C. y Fanelli A. (2008) Interpretability constraints for fuzzy information granulation. *Information Sciences* 178(24): 4585–4618.

[Nau03] Nauck D. (2003) Measuring interpretability in rule-based classification systems. En *Proceedings of the 12th IEEE International Conference on Fuzzy Systems*, volumen 1, páginas 196–201.

[PdO96] Pedrycz W. y de Oliveira J. (1996) Optimization of fuzzy models. *IEEE Trans. Systems Man Cybernet. Part B* 26(4): 627–636.

[PK08] Pulkkinen P. y Koivisto H. (2008) Fuzzy classifier identification using decision tree and multiobjective evolutionary algorithms. *International Journal of Approximate Reasoning* 48(2): 526–543.

[Ros67] Rosenberg R. (1967) *Simulation of genetic populations with biochemical properties.* PhD thesis, Univ. Michigan, Ann Harbor, Michigan.

[SBKvNL98] Setnes M., Babuška R., Kaymak U., y van Nauta Lemke H. R. (1998) Similarity measures in fuzzy rule base simplification. *IEEE Trans. Syst., Man, Cybern. - Part B: Cybernetics* 28(3): 376–386.

[Sch85] Schaffer J. (1985) Multiple objective optimization with vector evaluated genetic algorithms. En *Proc. First Int. Conf. on Genetic Algorithms*, páginas 93–100. Pittsburgh, USA.

[SD94] Srinivas N. y Deb K. (1994) Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary Computation* 2: 221–248.

[SY93] Sugeno M. y Yasukawa T. (1993) A fuzzy-logic-based approach to qualitative modeling. *IEEE Transactions on Fuzzy Systems* 1(1): 7–31.

[Thr91] Thrift P. R. (1991) Fuzzy logic synthesis with genetic algorithms. En Belew R. K. y Booker L. B. (Eds.) *ICGA*, páginas 509–513. Morgan Kaufmann.

[Zad65] Zadeh L. A. (1965) Fuzzy sets. *Information and Control* 8: 338–353.

[Zad73] Zadeh L. A. (1973) Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Syst., Man, Cybern.* 3: 28–44.

[Zad75] Zadeh L. A. (1975) The concept of a linguistic variable and its application to approximate reasoning, parts i, ii and iii. *Information Science* 8, 8 and 9: 199–249, 301–357 and 43–80.

[ZG08] Zhou S. M. y Gan J. Q. (2008) Low-level interpretability and high-level interpretability: A unified view of data-driven interpretable fuzzy system modelling. *Fuzzy Sets and Systems* 159(23): 3091–3131.

[ZLT01] Zitzler E., Laumanns M., y Thiele L. (2001) Spea2: Improving the strength pareto evolutionary algorithm for multiobjective optimization. En *Proc. Evolutionary Methods for Design, Optimization and Control with App. to Industrial Problems*, páginas 95–100. Barcelona, Spain.

[ZT99] Zitzler E. y Thiele L. (1999) Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. *IEEE Transactions on Evolutionary Computation* 3(4): 257–271.