

Representación del conocimiento mediante mapas cognitivos difusos y conjuntos de términos lingüísticos difusos dudosos en la biomedicina

[Knowledge representation using fuzzy cognitive maps and hesitant fuzzy linguistic term sets]

Salah Hasan Saleh¹, Sol David Lopezdomínguez Rivas², Andrea Mariel Malave Gomez², Fawaz Saleem Mohsen³, and Maikel Leyva Vázquez⁴

¹Universidad de las Ciencias Informáticas, La Habana, Cuba

²Facultad de Ciencias Químicas, Universidad de Guayaquil, Guayaquil, Guayas, Ecuador

³Servicio de referencia nacional de cardiopatía y embarazo, La Habana, Cuba

⁴Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, Guayaquil, Guayas, Ecuador

Copyright © 2016 ISSR Journals. This is an open access article distributed under the *Creative Commons Attribution License*, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT: Fuzzy cognitive maps have received increasing attention for the representation of causal knowledge, being especially useful in biomedicine. Recently some extension for using fuzzy cognitive maps using the paradigm of computing with words in order to provide causal models that are easily understood have been proposed. Nevertheless, there are situations in which experts hesitate between several values to assess the causal relation. To this end, we propose the use of of hesitant fuzzy linguistic term sets. Finally, the paper presents an illustrative example of the model for biomedical knowledge representation.

KEYWORDS: fuzzy cognitive maps, computing with words, hesitant fuzzy linguistic term sets, biomedicine.

RESUMEN: Los mapas cognitivos difusos han recibido una creciente atención para la representación del conocimiento causal, siendo de especial utilidad para la representación del conocimiento en el biomedicina. Recientemente se han presentado extensiones para el trabajo con la computación con palabras. Sin embargo existen situaciones en las cuales los expertos dudan entre varios valores lingüísticos. En el presente trabajo se propone un modelo basado en mapas cognitivos difusos utilizando el paradigma de computación con palabras con el objetivo de proporcionar modelos causales que sean fácilmente comprensibles y conjuntos de términos lingüísticos difusos dudosos. Finalmente, el trabajo presenta un ejemplo ilustrativo del modelo presentado aplicado a la representación del conocimiento en la biomedicina.

PALABRAS CLAVE: mapas cognitivos difusos, computación con palabras, biomedicina, términos lingüísticos difusos dudosos.

1 INTRODUCCIÓN

Los Mapas Cognitivos Difusos (MCD) [1] han recibido una creciente atención para la representación de conocimiento. Los MCD son grafos causales que hacen uso de la lógica difusa brindando la posibilidad de representar retroalimentación y modelar la vaguedad propia de este tipo de relaciones [2].

Los mapas cognitivos fueron propuestos por Axelrod [3]. En estos, los nodos representan conceptos o variables de un dominio. Las conexiones indican la dirección de la causalidad junto al signo asociado que puede ser positivo (incremento causal) o negativo (decremento causal). Sin embargo en el mundo cotidiano los enlaces entre causa y efecto son frecuentemente imprecisos por naturaleza, existiendo distintos grados de causalidad, por lo que se definieron los MCD que hacen uso de la lógica difusa [4], ya que ofrece un marco adecuado para tratar con la causalidad imperfecta en los mapas cognitivos.

La lógica difusa permite expresar el grado de causalidad entre conceptos a través del empleo de valores borrosos en el intervalo $[-1,1]$ y mediante el uso de expresiones lingüísticas como "negativamente fuerte", "positivamente fuerte", "negativamente débil", "positivamente débil", etc. En estos casos, la información lingüística modela de forma flexible el conocimiento e implica procesos de computación con palabras (CWW) [5].

Aunque se han realizado algunos esfuerzos en el empleo de la computación con palabras (CWW) en MCD como es el caso del análisis estático [6] y el desarrollo de un modelo utilizando 2-tuplas lingüísticas para el análisis del diseño conceptual [7] y como ayuda a la toma de decisiones [8].

En el presente trabajo un modelo de mapas cognitivos difusos que utilizan el paradigma de computación con palabras con el objetivo de proporcionar modelos causales que sean fácilmente comprensibles y conjuntos de términos lingüísticos difusos dudosos. Adecuado para situaciones en las cuales los expertos dudan entre varios valores lingüísticos

Este trabajo se estructura del siguiente modo: En la Sección 2 se abordan las temáticas relacionadas con los MCD y su empleo en la toma de decisiones. En la Sección 3 se presenta el modelo de representación basado en la 2-tuplas lingüística y conjuntos de términos lingüísticos difusos dudosos. En la Sección 4 se presenta el modelo propuesto y a continuación en la Sección 5 se introduce un caso de estudio aplicado a la representación del conocimiento a la biomedicina. El trabajo finaliza con las conclusiones y trabajos futuros.

2 MAPAS COGNITIVOS DIFUSOS Y COMPUTACIÓN CON PALABRAS

Los MCD fueron introducidos por Kosko [1] como una extensión de los mapas cognitivos. Los MCD describen la fortaleza de la relación mediante el empleo de valores borrosos en el intervalo $[-1,1]$. Para ello, los nodos representan conceptos causales y pueden modelar eventos, acciones, valores, metas o procesos. Así, los nodos constituyen una estructura de grafo difuso con retroalimentación para representar causalidad [9] y ofrecen un marco de trabajo más potente y flexible para representar el conocimiento humano para el razonamiento frente a los sistemas expertos tradicionales [10].

Dada la gran utilidad de los MCD, estos han sido extendidos para modelar diversas situaciones. Así, encontramos extensiones basadas en la teoría de los sistemas grises [11], intervalos [12], lógica difusa intuicionista [13], entre otras extensiones.

Los mapas cognitivos han presentado diversas aplicaciones para la representación del conocimiento en la medicina. Entre ellos se destacan evaluación del riesgo de desarrollar de cáncer de mamas [14], cáncer de próstata [15] y ayuda a la toma de decisiones en la radioterapia [16].

Stylios y otros [17] proponen la aplicación de los MCD a la toma de decisiones en la medicina, y denominan su modelo mapa cognitivo difuso competitivo. En él se introducen dos tipos de nodos fundamentales, los de decisión y los asociados a factores (figura 1) Se encuentran ejemplos de su aplicación en el diagnóstico de la patologías del lenguaje [17] y en la obstetricia [18]. Sin embargo la necesidad de que la relación entre los nodos sea numérica limita la interpretabilidad del moldeo.

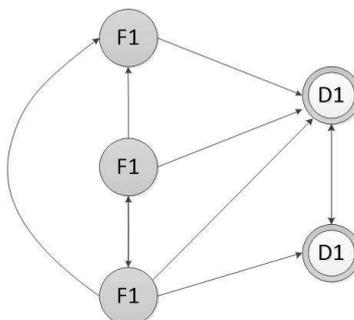


Fig. 1. Mapa cognitivo difuso competitivo constituido por dos alternativas (nodos de decisión) y tres criterios (factores)[19]

Recientemente se ha propuesto una extensión a los mapas cognitivos difusos competitivos utilizando computación con palabras y el modelo de las 2-tuplas lingüísticas denominado mapa cognitivo difuso competitivo lingüístico (LCFCM) [20]. La principal ventaja del modelo propuesto para el diagnóstico médico basado en mapas cognitivos difusos es que permite aumentar la interpretabilidad de los modelos causales, siendo este hecho de utilidad en el diagnóstico médico, sin embargo este modelo no permite el trabajo en situaciones frecuente en que los expertos dudan entre varios valores lingüísticos.

3 CONJUNTOS DE TÉRMINOS LINGÜÍSTICOS DIFUSOS DUDOSOS

La CWW es una metodología que permite realizar un proceso de computación y razonamiento utilizando palabras pertenecientes a un lenguaje en lugar de números. Dicha metodología permite crear y enriquecer modelos de decisión en los cuales la información vaga e imprecisa [21] es representada a través de variables lingüísticas.

El modelo de representación lingüística de 2-tuplas permite realizar procesos de computación con palabras sin pérdida de información, basándose en el concepto de traslación simbólica.

Sea $S = \{s_0, s_1, \dots, s_g\}$ un conjunto de términos lingüísticos y $\beta \in [0, g]$ un valor en el intervalo de granularidad de S.

Definición 1: [22] La Traslación Simbólica de un término lingüístico, s_i , es un número valorado en el intervalo $[-.5, .5]$ que expresa la diferencia de información entre una cantidad de información expresada por el valor $\beta \in [0, g]$, obtenido en una operación simbólica y el valor entero más próximo, $i \in \{0, \dots, g\}$ que indica el índice de la etiqueta lingüística (s_i) más cercana en S.

A partir de este concepto se desarrolló un nuevo modelo de representación de la información lingüística el cual hace uso de un par de valores o 2-tuplas. Este modelo de representación define un conjunto de funciones que facilitan las operaciones sobre 2-tuplas.

Definición 2: [22] Sea $S = \{s_0, s_1, \dots, s_g\}$ un conjunto de términos lingüísticos y $\beta \in [0, g]$ un valor que representa el resultado de una operación simbólica, entonces la 2-tupla lingüística que expresa la información equivalente a β , se obtiene usando la siguiente función:

$$\Delta: [0, g] \rightarrow S \times [-.5, .5]$$

$$\Delta(\beta) = (s_i, \alpha), \text{ con } \begin{cases} s_i, & i = \text{round}(\beta) \\ \alpha = \beta - i, & \alpha \in [-.5, .5] \end{cases} \quad (1)$$

Donde round es el operador usual de redondeo, s_i , es la etiqueta con índice más cercano a β y α es el valor de la traslación simbólica.

Cabe señalar que $\Delta^{-1}: \langle S \rangle \rightarrow [0, g]$ es definida como $\Delta^{-1}(s_i, \alpha) = i + \alpha$. De este modo, una 2-tupla lingüística $\langle S \rangle$ queda identificada con su valor numérico en $[0, g]$.

Recientemente se han empleado los conjuntos difusos dudosos HFS-del inglés Hesitant Fuzzy Set [23]. Estos conjuntos hesitant han sido aplicados a la definición de conjunto de términos lingüísticos [24] permitiendo el uso de gramáticas libres de contexto permitiendo la creación de expresiones lingüísticas en las evaluaciones de los expertos, haciendo de este modo más flexible el proporcionar información relacionada con el problema de toma de decisiones [25].

Sea $S^g = \{s_0, \dots, s_g\}$ el conjunto ordenado finito de $g+1$ términos, o valores posibles de una variable lingüística ϑ . Sea HS un subconjunto ordenado finito de términos de S^g . Según [9], la gramática libre de contexto GH que produce expresiones gramaticales ll resultantes en H_s mediante la función E_{GH} , nos permite tener expresiones lingüísticas para evaluar alternativas.

Por ejemplo, para $N = 4$ y $S = \{nulo, bajo, medio, alto, completo\}$ las valoraciones del experto se transformarían en conjunto de términos lingüísticos (Hesitant Fuzzy Linguistic Term Set - HFLTS) de la siguiente manera [24]:

$y_1 =$ inferior a bajo, siendo ll_1 : menor o igual que s_1 y por tanto $HS = \{s_0, s_1\}$.

$y_2 =$ superior a la media, siendo ll_2 : mayor o igual que s_3 y por tanto $HS = \{s_3, s_4\}$.

$y_3 =$ normal, siendo ll_3 : entre s_2 y s_2 y por tanto $HS = \{s_2, s_3, s_3\}$.

$y_4 =$ media, siendo $ll_4 = s_2$ y por tanto $HS = \{s_2\}$.

4 MODELO PROPUESTO

En esta sección se presenta el procedimiento empleado para obtención del modelo de diagnóstico de enfermedades cardiovasculares basado en computación con palabras y mapas cognitivos difusos competitivos. El mismo tiene como propósito fundamental mejorar la interpretabilidad de los modelos de diagnósticos de este tipo de enfermedades. Las fases incluidas en el modelo propuesto se representan gráficamente en la Figura 2 y son detalladas a continuación:



Fig. 2. Fases del modelo propuesto

4.1 REALIZAR PREPARACIÓN PREVIA

En primer lugar se identifican las fuentes de información a incluir en el estudio, de ser posible recomendándose la participación de múltiples expertos que representan distintos puntos de vista del sistema a modelar. Adicionalmente se definen conjunto (s) de términos lingüísticos que serán empleados para el modelado de las relaciones. Siendo:

$S = \{s_0, \dots, s_g\}$ el conjunto de términos lingüísticos.

Para que una fuente de información pueda expresar con mayor facilidad el conocimiento es necesario que disponga de un conjunto apropiado de descriptores lingüísticos. Con el propósito de facilitar el trabajo se recomienda el cumplimiento de las siguientes reglas [26]:

- El número de términos a la izquierda y a la derecha del término medio debe ser el mismo.
- La cardinalidad de un conjunto de términos lingüísticos no debe ser demasiado pequeña como para imponer una restricción de precisión a la información que quiere expresar cada fuente de información, y debe ser lo suficientemente grande como para permitir hacer una discriminación de las valoraciones en un número limitado de grados.

4.2 DEFINIR SÍNTOMAS, SIGNOS Y ENFERMEDADES

Se determinan los síntomas, signos y enfermedades, los cuales pasan a ser nodos en el modelo. En el modelo propuesto modelo se clasifican los nodos de dos formas, los síntomas y nodos son factores y las enfermedades son nodos de decisión. Esta clasificación permite modelar el problema siguiendo el concepto de los mapas cognitivos difusos competitivos.

4.3 DETERMINAR RELACIONES EXISTENTES

Cada experto expresa las relaciones haciendo uso de los términos lingüísticos definidos. El peso de la conexión que va del concepto C_i al concepto C_j , se representa mediante se emplearan conjunto lingüísticos difusos dudosos. Se hace usos del conjunto de etiquetas lingüísticas se hace uso de expresiones lingüísticas representadas mediante HFLTS.

5 RESULTADOS Y DISCUSIÓN

A continuación, para ilustrar la aplicación práctica del modelo propuesto, se muestra un ejemplo ilustrativo a la medicina en el que se muestra la relación de distintos síntomas (N_1, \dots, N_5) con una enfermedad (N_6). La descripción de los nodos aparece en la Tabla 1.

Tabla 1. Conceptos utilizados

Síntomas/Signos	Concepto	Descripción
N_1	Palpitaciones cardíacas	<ul style="list-style-type: none"> Sensaciones de latidos cardíacos que se perciben como si el corazón estuviera latiendo con violencia o acelerando.
N_2	Presión arterial	<ul style="list-style-type: none"> Es la fuerza o presión que ejerce la sangre contra las paredes de las arterias, llevando de esa forma la sangre a todas las partes del cuerpo.
N_3	Cefalea	<ul style="list-style-type: none"> Hace referencia a los dolores y molestias localizadas en cualquier parte de la cabeza.
N_4	Visión borrosa	<ul style="list-style-type: none"> Incapacidad del sistema de visión para percibir, detectar o identificar objetos especiales con unas condiciones de iluminación buenas.
N_5	Confusión mental	<ul style="list-style-type: none"> Disminución de la actividad de la conciencia. Existen varios grados que van desde una leve obnubilación hasta el estado de estupor.
N_6	Hipertensión arterial	<ul style="list-style-type: none"> Una enfermedad crónica caracterizada por un incremento continuo agudo o crónico de las cifras de la presión arterial.

Para este ejemplo ilustrativo se propone el conjunto de términos lingüísticos mostrado en la Figura 2 para representar las relaciones causales.

s_0	Cero (C)	(0.0,0.0,0.25)
s_1	Positivamente débil (PD)	(0.0,0.25,0.50)
s_2	Positivamente media (PM)	(0.25,0.50,0.75)
s_3	Positivamente fuerte (PF)	(0.50,0.75,1)
s_4	Positivamente muy fuerte (PMF)	(0.75,1,1)

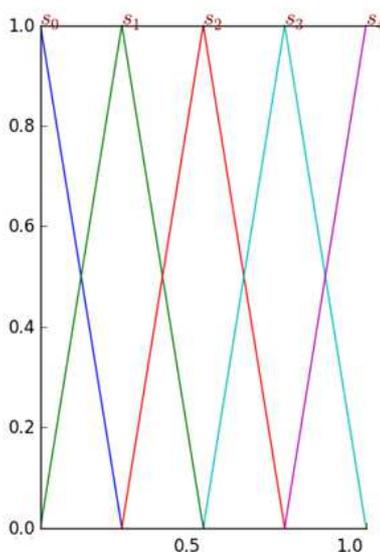


Fig. 3. Conjunto de etiquetas lingüísticas

Se obtuvo la siguiente matriz de adyacencia a partir de las valoraciones.

Tabla 2. Matriz de Adyacencia

C	C	C	C	C	C
Entre PM y PMF	C	C	C	C	Mayor que PM
C	PMF	C	C	C	Entre PM y PMF
C	C	C	C	C	Menor que PM
C	C	C	C	C	PD
Mayor que PF	C	C	C	C	C

A continuación se muestra la matriz de adyacencia en H_s mediante la transformación de cada expresión lingüística en HFLTS.

Tabla 3. Matriz de adyacencia en H_s

C	C	C	C	C	C
{PM,PF,PMF}	C	C	C	C	{PM,PF,PMF}
C	PMF	C	C	C	{PM,PMF}
C	C	C	C	C	{C,PD,PM}
C	C	C	C	C	PD
{PF,PMF}	C	C	C	C	C

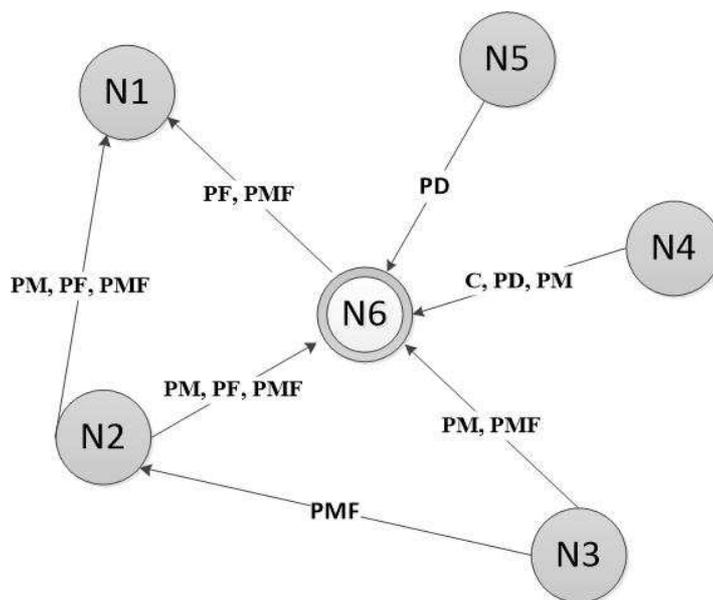


Fig. 4. MCDLD obtenido

La propuesta presentada muestra múltiples ventajas con respecto a la formulación original de los MCD y su aplicación al análisis de escenarios [27], logrando una mayor flexibilidad para que los expertos expresen sus valoraciones y una mayor interpretabilidad en el modelo al emplear términos lingüísticos. Al emplear el modelo de representación lingüístico basado en 2-tuplas se mejora los problemas de pérdida de información en los procesos CWW con respecto a otros modelos [22].

6 CONCLUSIONES

Los mapas cognitivos difusos se han demostrado como una herramienta útil de apoyo a la toma de decisiones. En este trabajo se ha presentado un nuevo modelo de mapas cognitivos basados lo conjunto lingüísticos dudosos. Además, se ha mostrado un ejemplo ilustrativo aplicado a la medicina.

La propuesta de esta contribución está centrada en aumentar la interpretación de los mapas cognitivos difusos y su facilidad de uso por los expertos, proporcionando resultados lingüísticos que sean fácilmente interpretables. Para ello, se propone que se represente la información a través de valores lingüísticos y se opere sobre ellos a través del modelo lingüístico basado en 2-tuplas. De este modo, los modelos mentales obtenidos son más cercanos al modo de pensar de los decisores.

Como trabajos futuros se plantea el desarrollo de un modelo de inferencia basado en los Mapas Cognitivos difusos intervalares [28]. El desarrollo de nuevos algoritmos de aprendizaje es otra área de trabajo.

REFERENCIAS

- [1] B. Kosko, "Fuzzy cognitive maps," *International Journal of Man-Machine Studies*, vol. 24, pp. 65-75, 1986.
- [2] M. Leyva-Vázquez, Karina Pérez-Teruel, A. Febles-Estrada, and J. Gulín-González, "Técnicas para la representación del conocimiento causal. Un estudio de caso en Informática Médica.," *ACIMED*, vol. 24, 2013.
- [3] R. M. Axelrod, *Structure of decision: The cognitive maps of political elites*: Princeton University Press Princeton, NJ, 1976.
- [4] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338-353, 1965.
- [5] L. Marti and F. Herrera, "An overview on the 2-tuple linguistic model for computing with words in decision making: extensions, applications and challenges," *Information Sciences*, vol. 207, pp. 1-18, 2012.
- [6] S. Samarasinghea and G. Strickert, "A New Method for Identifying the Central Nodes in Fuzzy Cognitive Maps using Consensus Centrality Measure," presented at the 19th International Congress on Modelling and Simulation, Perth, Australia, 2011.
- [7] A. Singh, "Architecture value mapping: using fuzzy cognitive maps as a reasoning mechanism for multi-criteria conceptual design evaluation," PhD Thesis, Missouri University of Science and Technology, Missouri, 2011.
- [8] K. Pérez-Teruel, M. Leyva-Vázquez, M. Espinilla, and V. Estrada-Sentí, "Computación con palabras en la toma de decisiones mediante mapas cognitivos difusos," *Revista Cubana de Ciencias Informáticas*, vol. 8, pp. 19-34, 2014.
- [9] C. W. Ping, "A Methodology for Constructing Causal Knowledge Model from Fuzzy Cognitive Map to Bayesian Belief Network," PhD Thesis, Department of Computer Science, Chonnam National University, 2009.
- [10] E. I. Papageorgiou, "Learning Algorithms for Fuzzy Cognitive Maps---A Review Study," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. PP, pp. 1-14, 2011.
- [11] J. L. Salmeron, "Modelling grey uncertainty with Fuzzy Grey Cognitive Maps," *Expert Systems with Applications*, vol. 37, pp. 7581-7588, 2010.
- [12] E. Papageorgiou, C. Stylios, and P. Groumpos, "Introducing Interval Analysis in Fuzzy Cognitive Map Framework Advances in Artificial Intelligence." vol. 3955, G. Antoniou, G. Potamias, C. Spyropoulos, and D. Plexousakis, Eds., ed: Springer Berlin / Heidelberg, 2006, pp. 571-575.
- [13] D. K. Iakovidis and E. Papageorgiou, "Intuitionistic Fuzzy Cognitive Maps for Medical Decision Making," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 15, pp. 100-107, 2011.
- [14] J. Subramanian, A. Karmegam, E. Papageorgiou, N. Papandrianos, and A. Vasukie, "An integrated breast cancer risk assessment and management model based on fuzzy cognitive maps," *Computer methods and programs in biomedicine*, vol. 118, pp. 280-297, 2015.
- [15] W. Froelich, E. I. Papageorgiou, M. Samarinas, and K. Skriapas, "Application of evolutionary fuzzy cognitive maps to the long-term prediction of prostate cancer," *Applied Soft Computing*, vol. 12, pp. 3810-3817, 2012.
- [16] E. Papageorgiou, C. Stylios, and P. Groumpos, "The soft computing technique of fuzzy cognitive maps for decision making in radiotherapy," *Intelligent and Adaptive Systems in Medicine*, pp. 106-127, 2008.
- [17] Stylios, V. C. Georgopoulos, G. A. Malandraki, and S. Chouliara, "Fuzzy cognitive map architectures for medical decision support systems," *Applied Soft Computing*, vol. 8, pp. 1243-1251, 2008.
- [18] C. S. Stylios and V. C. Georgopoulos, "Fuzzy Cognitive Maps for Medical Decision Support; A paradigm from obstetrics," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, 2010, pp. 1174-1177.

- [19] M. Leyva-Vázquez, "Modelo de Ayuda a la Toma de Decisiones Basado en Mapas Cognitivos Difusos," Doctor en Ciencias Técnicas, UCI. Doctor en Ciencias Técnicas, La Habana, 2013.
- [20] S. H. Saleh, M. Leyva-Vázquez, J. P. F. Rodríguez, and F. S. Mohsen, "Modelo para la representación de las interrelaciones entre síntomas, signos y enfermedades cardiovasculares basadas en computación con palabras y mapas cognitivos difusos competitivos," *Investigación Tecnología e Innovación*, vol. 6, 2015.
- [21] F. Herrera, S. Alonso, F. Chiclana, and E. Herrera-Viedma, "Computing with words in decision making: foundations, trends and prospects," *Fuzzy Optimization and Decision Making*, vol. 8, pp. 337-364, 2009.
- [22] F. Herrera and L. Martínez, "A 2-tuple fuzzy linguistic representation model for computing with words," *Fuzzy Systems, IEEE Transactions on*, vol. 8, pp. 746-752, 2000.
- [23] V. Torra and Y. Narukawa, "On hesitant fuzzy sets and decision," in *Fuzzy Systems, 2009. FUZZ-IEEE 2009. IEEE International Conference on*, 2009, pp. 1378-1382.
- [24] R. M. Rodriguez, L. Martinez, and F. Herrera, "Hesitant fuzzy linguistic term sets for decision making," *Fuzzy Systems, IEEE Transactions on*, vol. 20, pp. 109-119, 2012.
- [25] R. Montes, A. M. Sanchez, P. Villar, and F. Herrera, "Aplicación del uso de valoraciones hesitant lingüísticas para una consultora de inmuebles."
- [26] M. Espinilla, "Nuevos Modelos de Evaluación Sensorial con Información Lingüística," *SCHOOL= Universidad de Jaén, YEAR= 2009*, 2009.
- [27] M. Y. Leyva Vázquez, K. Y. Pérez Teurel, A. Febles Estrada, and J. Gulín González, "Modelo para el análisis de escenarios basados en mapas cognitivos difusos: estudio de caso en software biomédico," *Ingeniería y Universidad*, vol. 17, pp. 375-390, 2013.
- [28] E. Papageorgiou, C. Stylios, and P. Groumpos, "Introducing interval analysis in fuzzy cognitive map framework," in *Advances in Artificial Intelligence*, ed: Springer, 2006, pp. 571-575.