

Tipo de artículo: Artículo original
Temática: Inteligencia Artificial
Recibido: 7/05/2014 | Aceptado: 23/05/2014

La toma de decisiones en los Sistemas Tutoriales Inteligentes utilizando el agrupamiento conceptual

Decision making in intelligent tutoring systems using conceptual clustering

Yunia Reyes González^{1*}, Dra. Natalia Martínez Sánchez¹

¹Departamento de Programación de la Facultad 2. Universidad de las Ciencias Informáticas, Carretera a San Antonio de los Baños, km 2 ½, Torrens, Boyeros, La Habana, Cuba. CP.: 19370. Correo-e: natalia@uci.cu

*Autor para correspondencia: yrglez@uci.cu

Resumen

Las técnicas de Inteligencia Artificial son válidas para enfrentar la construcción de un Sistema Tutorial Inteligente dado por sus aspectos afines. Éstas utilizan conocimiento sobre un dominio específico para inferir una solución similar a la obtenida por una persona experimentada en el dominio del problema. Por su parte los Sistemas Tutoriales Inteligentes utilizan la información almacenada sobre las características del estudiante para adaptar el proceso de enseñanza-aprendizaje del mismo a la materia a enseñar. No todos los paradigmas de la Inteligencia Artificial facilitan la concepción de este tipo de sistemas de enseñanza-aprendizaje, donde lo fundamental para su desarrollo es determinar cómo representar el conocimiento requerido para sus módulos y a partir de dicho conocimiento realizar una prescripción del estudiante para que el sistema se adapte a sus características. Sin embargo, similitudes del modelado del estudiante en un Sistema Tutorial Inteligente y los algoritmos de agrupamiento conceptual son factores a estudiar para concebir un diagnóstico adecuado del qué y cómo enseñar dependiendo del estudiante.

Palabras clave: Agrupamiento conceptual, Modelado del estudiante, Sistema Tutorial Inteligente.

Abstract

The Artificial Intelligence techniques are valid to address the construction of an Intelligent Tutorial System given its related aspects. They use knowledge about a specific domain to infer a similar solution to that obtained by a person experienced in the domain of the problem. Meanwhile Intelligent Tutoring Systems using stored information about student characteristics for adapt the teaching-learning process thereof to the subject to be taught. Not all Artificial Intelligence paradigms facilitate the design of such systems of teaching and learning, where the main for their development is to determine how to represent the knowledge required for their modules and from that knowledge make a diagnostic of the student for that the system to adapt to your features. However, student modeling similarities in an Intelligent Tutorial System and conceptual clustering algorithms are factors to consider to design a proper diagnosis of what and how to teach depending on the student.

Keywords: *Conceptual clustering, Intelligent Tutorial System, Student modeling.*

Introducción

La toma de decisiones en un Sistema Tutorial Inteligente (STI) está estrechamente relacionada con el modelado del estudiante, problema central en el diseño y desarrollo de los STI. El proceso del modelado del estudiante consiste en inferir el estado cognitivo de éste utilizando la información almacenada en el modelo del estudiante (Ovalle, 2007).

Uno de los problemas centrales en muchas disciplinas lo constituye el análisis de los datos. Se ha desarrollado un conjunto grande de herramientas para la solución de este problema a partir de diferentes enfoques. En la medida en que los datos se hacen más complejos, por ejemplo, dejan de ser exclusivamente numéricos para presentarse mezclados con datos de naturaleza cualitativa, con subjetividad e imprecisión, aparecen descripciones de objetos incompletas (ausencia de información), mayor es la complejidad del análisis que se pretende realizar y mayor es la dificultad de extraer información útil de los mismos.

A finales de los años 70, Ryszard S. Michalski, introdujo un conjunto de ideas que han dado en llamarse agrupamiento conceptual. En este enfoque, lo que se propone es aportar una información adicional a la estructuración de un espacio; no se pretende sólo decir quiénes forman un agrupamiento sino, además, decir en alguna medida por qué, es decir, brindar más información acerca de los agrupamientos, caracterizarlos a partir de las propiedades que éstos cumplen, definidas sobre la base de los rasgos en términos de los cuales se describen a los objetos en estudio. El agrupamiento conceptual está compuesto de dos tareas fundamentales: el agrupamiento de entidades en el que se

determinan subconjuntos útiles de una muestra de objetos, y la caracterización, la cual determina un concepto para cada subconjunto descubierto. Estas dos tareas no tienen necesariamente que ser independientes ni realizarse en un orden determinado (Pons, 2002)

Los agrupamientos obtenidos se pueden describir, por tanto, de forma extensional e intencional. La descripción extensional se realiza mediante la enumeración de los objetos que los componen y la descripción intencional contiene las propiedades que caracterizan al agrupamiento (Ruiz., 2013).

En este trabajo se describe un método para el modelado del estudiante en un STI utilizando el enfoque del agrupamiento conceptual. Partiendo de un conjunto de modelos de estudiantes, descritos a través de rasgos que toman valores de diferentes tipos, se estructuran en diferentes clases, a las cuales se les asocia la descripción conceptual que las caracterizan. Esta descripción se corresponde con el modelado de los estudiantes que conforman la clase y a partir del cual se toma la decisión de qué material didáctico se adapta al estado cognitivo del alumno.

Metodología computacional

El desarrollo de un STI tiene implícito en mayor o menor medida la aplicación de técnicas de Inteligencia Artificial. Estos pueden construirse implementando un tipo de Sistema Basado en el Conocimiento, según resulte de la Ingeniería del Conocimiento que lleva implícito todo STI (Martínez, 2010).

Es precisamente en el proceso de IC donde están centradas las ideas fundamentales de este trabajo, basadas en aplicar algoritmos del enfoque lógico-combinatorio para la construcción del modelo del estudiante y específicamente algoritmos de agrupamiento conceptuales en el modelado del estudiante.

Los algoritmos de agrupamiento conceptual se pueden clasificar en dos grandes grupos:

- 1) **Algoritmos Incrementales:** Basan su funcionamiento en la adaptación de los agrupamientos (o conceptos) con los nuevos objetos que se le van presentando, es decir, cada vez que se analiza un nuevo objeto este se clasifica, mediante una cierta estrategia, en uno de los agrupamientos ya existentes o se crean nuevos grupos. Por lo general, estos algoritmos se han desarrollado para los casos en que el conjunto de objetos a estructurar

no está completamente dado, es decir, es dinámico. Ejemplos de estos algoritmos son: UNIMEM (Lebowitz, 1987), COBWEB (Fisher, 1987) y Galois(Carpineto & Romano 1996), entre otros.

- 2) **Algoritmos No incrementales:** Estructuran una muestra de objetos sin presuponer que estos se estructuran de uno en uno. Ejemplo: CLUSTER/PAF (Michalski, 1979), CLUSTER/2(Michalski, 1983), CLUSTER/S(Stepp & Michalski, 1986), WITT(Hanson & Bauer, 1989) y K-Means conceptual(Ralambondrainy, 1995), LCconceptual(Martínez-Trinidad, & otros, 2001) y RGC(Pons-Porrata, & otros, 2002), entre otros.

El modelo que se propone en esta investigación retoma las ideas básicas de los algoritmos conceptuales no incrementales que se corresponden con el enfoque lógico combinatorio. Este enfoque encuentra su basamento en ramas como: la lógica matemática, la teoría de testores y la teoría clásica de conjuntos, las cuales constituyen el sustento matemático del procedimiento que se describe.

Modelo para elaborar un Sistema Tutorial Inteligente utilizando el agrupamiento conceptual.

Un STI lo componen tres módulos fundamentales. El Módulo del Estudiante que almacena la información relacionada con el alumno, a través de él se determina ¿Qué conoce el estudiante? y a partir de la respuesta a esta interrogante se infiere ¿Qué enseñar? y ¿Cómo enseñar?, informaciones representadas en el Módulo del Dominio y Módulo Pedagógico respectivamente.

a. Estructura del Módulo del Estudiante.

Puede afirmarse que el modelo del estudiante es un problema de investigación que debe enfocarse desde todas sus aristas con el fin de obtener una representación de las características del estudiante completa y precisa. Algunos autores toman en consideración características tales como: el estilo de aprendizaje, el nivel de conocimiento, la información personal o la combinación de algunas de ellas.

En el modelo utilizando el agrupamiento conceptual que se describe en este trabajo el Módulo del Estudiante se representa a través de una matriz inicial (MI) como ilustra la Figura 1, donde n es la cantidad de rasgos (características que describen los modelos de estudiantes) y m el número de objetos (modelos de estudiantes).

	x_1, \dots, x_n
O_1	$x_1(O_1), \dots, x_n(O_1)$
\vdots	\vdots
O_m	$x_1(O_m), \dots, x_n(O_m)$

Figura 1. Matriz Inicial

b. Módulo del Dominio y Módulo Pedagógico

El Módulo del Dominio y Módulo Pedagógico están estrechamente relacionados y se representan a través de los materiales didácticos que se elaboran según los estados cognitivos de los estudiantes, facilitando al estudiante un proceso de enseñanza aprendizaje personalizado.

Módulo del Dominio

El módulo del dominio, denominado también por muchos autores como módulo experto, proporciona los conocimientos del dominio. Satisface dos propósitos diferentes. En primer lugar, presentar la materia de la forma adecuada para que el alumno adquiriera las habilidades y conceptos, lo que incluye la capacidad de generar preguntas, explicaciones, respuestas y tareas para el alumno. En segundo lugar, el módulo del dominio debe ser capaz de resolver los problemas generados, corregir las soluciones presentadas y aceptar aquellas soluciones válidas que han sido obtenidas por medios distintos.

En este módulo, el conocimiento a ser enseñado por el STI debe organizarse pedagógicamente para facilitar el proceso de enseñanza.

Módulo Pedagógico

Decide qué, cómo y cuándo enseñar los contenidos del dominio, adaptando sus decisiones pedagógicas a las necesidades del estudiante. Algunos autores le denominan módulo tutor, ya que es el encargado de comparar las características de los estudiantes con el contenido a enseñar y elegir la mejor forma de tomar las decisiones pedagógicas oportunas, adaptándose en cada momento al estudiante.

La figura 2 ilustra un ejemplo de un modelado del estudiante. A partir de los valores que toman los rasgos o características que describen el modelo del estudiante, se le asigna uno u otro material didáctico que permita un proceso de enseñanza-aprendizaje adaptativo.

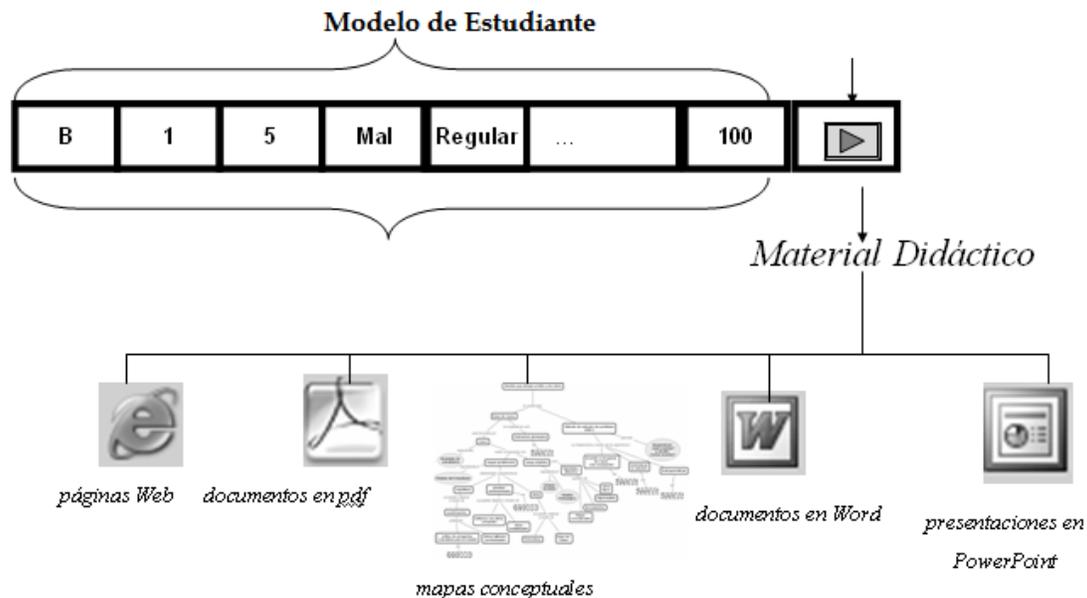


Figura 2. Ejemplo del modelado de un estudiante.

Modelado del estudiante utilizando el agrupamiento conceptual

Siguiendo la idea general expresada en (Martínez, 2012), la elaboración de un STI está determinado por tres etapas fundamentales, estrechamente relacionadas y con un orden de precedencia establecido, que facilitan definir los modelos de estudiantes y materiales didácticos a utilizar quedando así definido los tres módulos fundamentales de este tipo de sistema de enseñanza-aprendizaje.

Si se está en presencia de un problema no supervisado, por qué elegir los algoritmos de agrupamiento conceptuales y no los algoritmos de agrupamiento tradicionales para este tipo de problemas como aparece en propuestas realizadas por los propios autores de este trabajo en publicaciones anteriores. La respuesta a esta interrogante está dada en que una limitación de estos métodos consiste en que no obtienen la interpretación conceptual de los agrupamientos formados. El problema de la interpretación de los agrupamientos obtenidos es dejado al especialista o analista de

datos. Esta desventaja es significativa, ya que el especialista para los fines de su investigación o labor requiere no sólo los agrupamientos, sino además las descripciones conceptuales de los mismos. Esto es necesario incluso cuando el número de objetos en los agrupamientos no es muy grande (varias decenas de descripciones de objetos en términos de unas 10 ó 20 variables), pues la interpretación manual de los agrupamientos se hace muy compleja.

La característica distintiva del agrupamiento conceptual son sus dos etapas fundamentales, la extensional e intencional. En el modelo que se propone en este trabajo los algoritmos 1 y 2 y el algoritmo 3 describen las etapas extensional e intencional respectivamente.

Algoritmo 1. Agrupamiento no supervisado (Ruiz Shulcloper, 1995)

Entrada: MI

Salida: MA (objetos agrupados).

P1: Construir la Matriz de Semejanza¹ utilizando una función de semejanza.

P2: Calcular el umbral de semejanza β_0 :

$$\beta_0 = \frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \beta(o_i, o_j)$$

donde m: número de casos, i y j recorren las filas y columnas respectivamente.

P3: Agrupar siguiendo un criterio agrupacional

La selección de rasgos es cuestión central tanto en la definición del modelo del estudiante como en el modelado del estudiante. Potencialmente, en el conjunto de modelos de estudiantes podrían estar todas las propiedades que los describen; pero existen rasgos inútiles que carecen de importancia de acuerdo al dominio de aplicación.

El conjunto de rasgos determina qué información será almacenada en memoria para cada uno de sus elementos, la cual debe permitir el posterior modelado del estudiante, dada la descripción de un nuevo estudiante. Otro aspecto es que dentro del conjunto de rasgos que se seleccionan no todos tienen la misma importancia y esta diferencia debe tenerse en cuenta para comparar objetos.

¹Matriz que se obtiene como resultado de comparar todos los objetos de la matriz de aprendizaje.

De forma general el problema de selección de rasgos consiste en encontrar el subconjunto de rasgos que mejor describe los objetos del dominio; usualmente este subconjunto se encuentra maximizando o minimizando una función objetivo.

Algoritmo 2. Selección de los rasgos relevantes y cálculo de su importancia.

(Ruiz Shulcloper, 1995)

Entrada: MA.

Salida: TT (Testor de mayor relevancia)

P1: Formar la Matriz de Diferencia (MD).

P2: Formar la Matriz Básica (MB).

P3: Seleccionar el conjunto de testores.

P4: Calcular el peso de los rasgos que aparecen en la familia de testores:

$$P(x_i) = \frac{T_i}{|T|} \quad L(x_i) = \frac{\sum_{t \in T} \frac{1}{|T_i|}}{|T|}, \text{ donde}$$

T_i : número de testores donde aparece la pregunta (rasgo) i .

$|T|$: número de testores.

$|T_i|$: número de rasgos que forman el testor T_i

$$\varepsilon_i(x_i) = \alpha P_i(x_i) + \beta L_i(x_i),$$

donde $\alpha > 0$, $\beta > 0$ y $\alpha + \beta = 1$. α y β son dos parámetros que ponderan la participación o influencia de P_i (frecuencia de aparición) y L_i (longitud de los testores) respectivamente en ε . En el modelo se consideran $\alpha = \beta = 0.5$.

P5: Selección del conjunto de testores relevantes:

a. Para cada testor t_i calcular:

$$\psi_i(t_i) = \sum_{i=1}^{|t_i|} \varepsilon(x_i), \text{ en otras palabras, la magnitud calculada}$$

es la suma de la importancia de las preguntas que componen el testor.

- b. Seleccionar el testor de mayor $\psi_i(t)$, compuesto por los rasgos relevantes.

Seleccionado el testor de mayor relevancia utilizando el algoritmo 2, se forman los conceptos de los grupos obtenidos al aplicar el algoritmo 1.

El algoritmo 3 describe los conceptos de cada grupo obtenido después de aplicar el algoritmo 1. Los conceptos se describen a través de los valores que toman los rasgos que contiene el testor seleccionado. Cada rasgo está descrito por un par (x_i, μ_i) , donde x_i es el valor del rasgo i y μ_i es un valor de certeza.

Algoritmo 3. Descripción de los conceptos. (Pons, 2002)

Entrada: MA, TT

Salida: LC' //conjunto de los l-complejos para cada K_i

Paso 1. Para cada clase K_i , $i = 1, \dots, r$ se calcula la estrella

$G_r(K_i \setminus K_1, \dots, K_{i-1}, K_{i+1}, \dots, K_r)$ que estará formada por los $|TT|$ l-complejos [2].

Paso 2. Los conceptos que caracterizan a cada agrupamiento K_i serán los l-complejos obtenidos en el paso anterior.

Aplicando los anteriores algoritmos, el **Módulo del Estudiante** queda estructurado jerárquicamente siguiendo ideas de las diferentes jerarquías que han sido propuestas en modelos consultados en la literatura; pero no se utiliza ninguno de los enfoques tal y como han sido definidos en este trabajo.

Algoritmo 4. Modelado del estudiante

Entrada: MA y LC' (los l-complejos asociado a cada clase K_i)

O_t // nuevo modelo de estudiante

Salida: K_i seleccionada

Paso 1. Para cada K_i de MA calcular:

- a. $\beta_j(O_t, O_1)$ // donde O_1 son los objetos que se corresponden con los l-complejos. Es válido aclarar que sólo se comparan los rasgos que conforman el l-complejo.

b. Calcular:

$$\lambda_i(O_t) = \frac{\sum \beta(O_t, O_i)}{|I|} \quad // \text{ donde } O_t \text{ es el nuevo modelo de estudiante, } O_i \text{ los}$$

conceptos y $|I|$ el número de conceptos asociados a la clase i .

Paso 2.

- a. Seleccionar el $\max (\lambda_i(O_t))$
- b. Seleccionar la clase K_i // los objetos de la clase i se corresponden con los objetos más semejantes al nuevo modelo de estudiante.

Paso 3.

- a. Asignar al nuevo estudiante los materiales didácticos asociados a la clase seleccionada para lograr un proceso de enseñanza-aprendizaje personalizado.

El algoritmo 4 implementa el proceso del modelado del estudiante, que fundamentalmente se basa en dado un nuevo modelo de estudiante, a partir de las características que lo describe se diagnostica el estado cognitivo y se selecciona qué parte del dominio debe estudiar y qué estrategias pedagógicas seguir.

Resultados y discusión

Como resultado de este trabajo se conceptualizó un modelo jerárquico que utiliza criterios agrupacionales del enfoque lógico combinatorio sin aprendizaje, no referenciado en la literatura consultada con esta finalidad. En éste, se organizan los objetos en conjuntos sobre la base de criterios de semejanza definidos. Se identifican los objetos en diferentes agrupaciones, donde estas se generan de manera “natural” según el comportamiento global o particular de las semejanzas entre los objetos o atendiendo al cumplimiento de una cierta propiedad. No se necesita tener a priori ninguna información al respecto de las agrupaciones.

Para cada grupo de objetos (modelos de estudiantes) se determina el concepto que describe el grupo (algoritmos 2 y 3). La importancia de la utilización de algoritmos conceptuales para el modelado del estudiante en los STI radica en que el profesor puede contar con una descripción cualitativa de los diferentes modelados de estudiante dado los modelos de estudiantes.

Conclusiones

En este artículo se describe el resultado de una investigación que comenzó con un único objetivo de obtener una vía que facilitara el proceso de modelación del estudiante en los STI, el cual sobrepasó las expectativas de las autoras, ya que aporta a todo el proceso de implementación de un STI, desde la ingeniería del conocimiento necesaria hasta al método de inferencia que se utiliza en este tipo de sistema de enseñanza aprendizaje:

- La estructuración del Módulo del Estudiante en grupos aplicando la idea de los algoritmos clásicos de clasificación no supervisada facilita el proceso del modelado del estudiante, pues permite un análisis de la tendencia o comportamiento cognitivo de los estudiantes en el área del saber que se desarrolla el STI.
- El cálculo de los testores permite seleccionar aquellos rasgos que cuando no se consideran para describir el modelo de estudiante a diagnosticar se deteriora el desempeño o la precisión del modelado.
- A partir de la alternativa del cálculo de los testores que ofrece el enfoque lógico combinatorio se conforman los conceptos que describen los grupos en la fase intencional de este enfoque. Conceptos que no son más que descripciones cualitativas que se corresponden con el modelado del estudiante a partir del modelo de estudiante presentado. Las cuales permiten a los profesores o expertos obtener un modelado del estudiante más adecuado y elaborar los materiales didácticos adaptados al estado cognitivo-afectivo del estudiante, característica distintiva de los STI.

Se puede concluir que el ciclo de vida de un STI y todos los módulos que lo conforman puede ser implementado de forma natural si se hace corresponder como un problema de clasificación no supervisada utilizando el enfoque del agrupamiento conceptual. El módulo del estudiante se representa a través de una matriz inicial, conformada por objetos que se corresponden con los modelos de estudiantes. El modelado del estudiante no es más que el proceso de clasificación de un nuevo modelo de estudiante en la clase que más se adecua a su estado cognitivo y la posterior asignación de los materiales didácticos de dicha clase para lograr un proceso de enseñanza- aprendizaje personalizado.

Referencias

- CARPINETO, C. & ROMANO, G. (1996). "A lattice conceptual clustering system and its application to browsing retrieval. *Machine Learning*, 1996. 24(2): p. 95-122."
- FISHER, D. H. (1987). "Knowledge Acquisition Via Incremental Conceptual Clustering. *Machine Learning*, 1987. 2(2): p.139-172."
- HANSON, S. J. & BAUER, M. (1989). "Conceptual Clustering, Categorization, and Polymorphy. *Machine Learning* 3: p. 343-372. Kluwer Academic Publishers.
- LEBOWITZ, M. (1987). "Experiments with Incremental Concept Formation: UNIMEM. *Machine Learning*, 1987. 2(2):p. 103-138."
- MARTÍNEZ SÁNCHEZ, N. (2010). Sistemas Basados en Casos & Sistemas de Enseñanza-Aprendizaje Inteligentes. IE Comunicaciones. Revista Iberoamericana de Informática Educativa. Número 11, Enero - Junio 2010 pp 27-39. ISSN: 1699-4574 © ADIE.
- MARTÍNEZ, N, LORENZO, MARÍA MATILDE Y HURTADO, JORGE. 2012. Model for Designing Intelligent Tutorials Systems using Conceptual Maps and Knowledge-Based Systems. *IEEE Latin America Transactions*, Vol. 10, No. 6., págs. 2301-2308.
- MARTÍNEZ-TRINIDAD, J. F. & OTROS. (2001). "LC: A Conceptual Clustering Algorithm, in *Proceedings of the Second International Workshop on Machine Learning and Data Mining in Pattern Recognition 2001*, Springer-Verlag. p. 117-127."
- MICHALSKI, R. S. & STEPP, R. E (1979). "Conceptual Clustering: A Theoretical Foundation and a Method for Partitioning Data into Conjunctive Concepts, 1979: In *Textes des exposes du Seminaire organise par 'Institute de Recherche d'Informatique et d'Automatique (IRIA)*". :p. pp. 254-294.
- MICHALSKI, R. S. & STEPP, R. E. (1981). Concept-based Clustering versus Numerical Taxonomy. Invited paper submitted to *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

- OVALLE, D. 2007. Análisis funcional de la estrategia de aprendizaje individualizado adaptativo. Proyecto de investigación - DIME - Vicerrectoría de Investigación. Modelo de sistema multiagente de cursos adaptativos integrados con ambientes colaborativos : s.n., 2007.
- PONS-PORRATA, A. 2002. " RGC: a new conceptual clustering algorithm for mixed incomplete data sets". 2002, In Mathematical and Computer Modelling. p. 1375-1385."
- PONS, A. 2003. LEX: Un nuevo algoritmo para el cálculo de los testores típicos. Revista Ciencias Matemáticas. Cuba.
- RALAMBONDRAINY, H. (1995). "A conceptual version of the K-means algorithm. Pattern Recogn. Lett." **16**(11): p.1147-1157.
- RUIZ-SHULCLOPER, JOSÉ. 2013. Reconocimiento lógico combinatorio de patrones: teoría y aplicaciones. S.l. : tesis en opción al grado científico de doctor en ciencias. Centro de Aplicaciones de Tecnologías de Avanzada., 2013.
- RUIZ SHULCLOPER, JOSÉ Y LAZO CORTÉS, MANUEL. 1995. Introducción al Reconocimiento de Patrones: enfoque lógico combinatorio. s.l. : Serie Verde No. 51, CINVESTAV-IPN. México., 1995.
- STEPP, R. E. & MICHALSKI, R. S. (1986). "Conceptual Clustering: Inventing Goal-Oriented Classifications of Structured Objects, in Machine Learning: An Artificial Intelligence Approach, J.G.C. In R. S. Michalski, and T. M. Mitchell Editor 1986, Morgan Kaufmann."