

DISEÑO DE UN SISTEMA BASADO EN AGENTES PARA APOYAR EL DIAGNÓSTICO DE LA CALIDAD DEL SEMEN HUMANO

ESMERALDA RAMOS¹, HAYDEMAR NÚÑEZ¹, ROBERTO CASAÑAS²

¹Laboratorio de Inteligencia Artificial, Centro de Ingeniería de Software y Sistemas, Facultad de Ciencias, Escuela de Computación, Universidad Central de Venezuela

{esmeralda.ramos, haydemar.nunez}@ciens.ucv.ve

²Cátedra de Física y Análisis Instrumental, Escuela de Bioanálisis,

Facultad de Medicina, Universidad Central de Venezuela

robertoc@med.ucv.ve

Recibido: junio de 2011

Recibido en forma final revisado: octubre de 2011

RESUMEN

El análisis de la calidad del semen humano es un complejo proceso de toma de decisiones entre múltiples clases espermáticas. Requiere del profesional que lo realice una amplia gama de conocimientos y gran experticia. Con la finalidad de apoyar el diagnóstico de la calidad del semen, en este trabajo se presenta el diseño de un sistema basado en agentes que estará en capacidad de: a) clasificar un eyaculado utilizando un mecanismo que combina razonamiento basado en casos y razonamiento basado en reglas con lógica difusa, b) gestionar historias médicas de los pacientes, c) recuperar información especializada desde una ontología y d) analizar la morfología de las cabezas de los espermatozoides a partir de imágenes digitalizadas, utilizando técnicas de procesamiento de imágenes y paradigmas de aprendizaje artificial (redes neuronales y máquinas de soporte vectorial). La ontología desarrollada, que estandariza y organiza el conocimiento del dominio del análisis del líquido seminal, constituye la primera versión disponible. El análisis del sistema se realizó con la metodología INGENIAS y se implementó con un enfoque orientado a la Web.

Palabras clave: Agentes, Ontologías, INGENIAS, Razonamiento basado en casos, Redes neuronales, Semen humano.

DESIGNING AN AGENT-BASED SYSTEM TO SUPPORT THE DIAGNOSIS OF HUMAN SEMEN QUALITY

ABSTRACT

The analysis of human semen quality is a complex process that requires professionals to have a wide range of knowledge and great expertise. In order to support them in this task, in this paper an design of an agent-based system is proposed. The agents are able to: a) Classify the ejaculate quality using a fuzzy hybrid model of case-based reasoning and rule-based reasoning, b) Manage patients' medical histories, c) Retrieve specialized information from a domain ontology of Andrology, and d) Analyze the morphology of sperm heads from digitized images using image processing techniques and machine learning techniques (neural networks and support vector machines). In addition, the ontology developed is the first version available for this domain. The system was design with the INGENIAS methodology using a Web-oriented approach.

Keywords: Agent, Ontologies, INGENIAS, Case-based reasoning, Neural networks, Human semen.

INTRODUCCIÓN

El concepto de agente provee una forma conveniente y eficaz de describir una entidad de software, que es capaz de actuar con cierto grado de autonomía para cumplir tareas en representación de las personas. Los Sistemas Basados en Agentes (SBA) son especialmente adecuados para resolver problemas difíciles que puedan dividirse en sub-problemas,

que usen conocimiento heterogéneo de naturaleza distribuida y que necesiten que las entidades involucradas en alcanzar la solución, se comuniquen y colaboren entre sí.

En el dominio médico, son muchos los problemas que presentan las características antes mencionadas, razón por la cual se ha observado un importante incremento en el desarrollo de SBA que apoyan y asesoran a los profesionales

en la resolución de problemas. Estos sistemas se están usando para diagnosticar y tratar enfermedades (Huang *et al.* 2007; Isern *et al.* 2007; Kumar & Syamala, 2007; Gupta & Pujari, 2009), monitorear pacientes (Corchado *et al.* 2008; Bajo *et al.* 2009; Su & Wu, 2009), gestionar recursos hospitalarios (Heine *et al.* 2004), entre otros.

Un problema fundamental en el dominio médico es el relacionado con el análisis de los líquidos biológicos; particularmente el Análisis del Líquido Seminal (ALS). Diagnosticar la calidad del semen es un proceso complicado de toma de decisiones entre múltiples opciones de clasificación, determinadas por las alteraciones de las variables estimadas durante el análisis.

Para realizar el ALS, los profesionales deben seguir rigurosamente las directrices que proporcionan los manuales de laboratorio, conocer los materiales e instrumentos que deben manipular, poseer un extenso conocimiento del dominio, tener capacidades para realizar de manera simultánea diversas pruebas de laboratorio y, además, deben manipular la información de las historias médicas de los pacientes. La combinación de estos elementos dificulta el proceso de análisis, razón por la cual un apoyo computacional sería bien recibido por los profesionales.

Dado que la utilización de SBA ha resultado conveniente en la resolución de problemas en dominios médicos, en este trabajo se propone el diseño de un sistema de este tipo para apoyar a los profesionales en el diagnóstico de la calidad del semen humano que permita, además, gestionar información de las historias médicas y proveer acceso a información específica del dominio. Para el análisis y diseño del sistema se utiliza la metodología INGENIAS (Pavón *et al.* 2005) y su implementación se lleva a cabo con un enfoque orientado a la Web.

SISTEMAS BASADOS EN AGENTES

Un agente es todo aquello que puede percibir el ambiente mediante sensores y responder en tal ambiente mediante actuadores (Russell & Norvig, 2004). Para Jennings & Wooldridge (2000), es un software que se encuentra encapsulado en un ambiente, que responde de manera oportuna y autónoma a los cambios que se producen en su entorno para satisfacer sus objetivos de diseño. Un tipo particular de agente son los Agentes de Información (AI), los cuales se caracterizan por proporcionar algunos de los siguientes servicios: adquisición de información, gestión de información (actualizan y mantienen repositorios de datos), recuperación de información, integración (combinan información heterogénea), observación y adaptación de la información (presentan información de manera adecuada) (Gandon *et al.* 2002).

Para desarrollar SBA, es necesario disponer de la tecnología apropiada lo que implica utilizar conceptos, teorías, métodos y herramientas. La tecnología de agentes contempla especificaciones técnicas y algorítmicas que abordan aspectos como interacciones, reacción y deliberación de las arquitecturas, aprendizaje, coordinación y comunicación (Luck *et al.* 2005). También deben considerarse las tecnologías de desarrollo como esquemas para descripciones estructurales y semánticas del conocimiento (ontologías y bases de conocimiento), lenguajes de programación, metodologías de desarrollo y plataformas de construcción.

Las metodologías de desarrollo de SBA proporcionan directrices para construir sistemas siguiendo un proceso de desarrollo que proporciona herramientas de soporte. Destacan metodologías como: Vowel Engineering, GAIA, INGENIAS, MAS-CommonKADS, Prometheus (Henderson-Seller & Giorgini, 2005), entre otras. Resulta de especial interés INGENIAS, porque permite la construcción de SBA de manera incremental, asegurando la corrección durante el desarrollo y considerando los estados del ciclo de vida del software; proporciona un lenguaje visual de definición para describir las entidades del sistema y las relaciones entre éstas, además de herramientas que soportan el proceso de desarrollo, específicamente el INGENIAS Development Kit-IDK- (<http://ingenias.sourceforge.net>). La unión de estos elementos facilita el entendimiento y da cuenta de manera gráfica del comportamiento, responsabilidades y recursos de los agentes.

ANÁLISIS DEL SEMEN HUMANO

La evaluación de la fertilidad masculina se basa en una serie de ensayos que proporcionan información de la calidad del eyaculado y de la competencia funcional del espermatozoide. La evaluación más utilizada en las consultas por infertilidad y chequeos pre y post operatorios, es el ALS o espermograma. El ALS consiste de un conjunto de pruebas macroscópicas y microscópicas que valoran cualitativa y cuantitativamente los espermatozoides (Mortimer, 1994).

Diagnosticar la calidad del semen es un proceso complicado de toma de decisiones entre múltiples opciones de clasificaciones espermáticas, determinadas a partir de las alteraciones de las variables estimadas durante las pruebas y por los grados de severidad (leve, moderado, severa) que éstas puedan asumir. Los profesionales usan la nomenclatura de la Tabla 1, la cual se basa en criterios del manual de la Organización Mundial de la Salud -OMS- (OMS, 2001) y en su experiencia.

Si se conjugan las alteraciones de las variables que

determinan las clases espermáticas, se pueden obtener hasta 51 combinaciones. A su vez, las clasificaciones cuyas alteraciones presentan grado de severidad (con * en la Tabla 1), pueden ser sub-clasificadas entre 3 y 81 subclases. Además, el profesional realiza recomendaciones y sugerencias sobre la base de variables que, aun cuando no intervienen en la determinación de la clasificación, complementan el diagnóstico.

Tabla 1. Nomenclatura para clasificar un eyaculado

Clasificación	Descripción
Normozoospermia	Eyaculado normal según valores de referencia
Oligozoospermia *	Concentración de espermatozoides menor al valor de referencia
Astenozoospermia *	Motilidad inferior al valor de referencia
Teratozoospermia	Cantidad de espermatozoides con formas normales, inferior al valor de referencia
Hipospermia *	Volumen menor al valor de referencia
Necrozoospermia *	Vitalidad inferior al valor de referencia
Azoospermia	Ausencia de espermatozoides
Oligo Asteno Terato Hipos Necrozoospermia*	Perturbación significativa de las cinco variables (se utilizan también combinaciones de 2, 3, y 4 variables)

(* Alteración con grado de severidad)

En la Tabla 2 se muestran los resultados reportados para una muestra de semen, cuya clasificación espermática y recomendación complementaria del diagnóstico sería la

siguiente:

“Oligozoospermia Leve - Astenozoospermia Moderada. La viscosidad aumentada dificulta el movimiento progresivo de los espermatozoides, el aumento puede deberse a infecciones”

Tabla 2. Resultados de las pruebas realizadas a una muestra de semen

Pruebas Macroscópicas	Pruebas Microscópicas
Volumen: 3.0 ml	Concentración: 16.5 millones / ml
Aspecto: Homogéneo	Formas Normales: 15%
pH: 7.8	Formas Anormales: 85%
Viscosidad: Aumentada	Inmóviles Vivos: 46%
Color: Blanquecino	Progresivos lentos: 17%
Licuefacción: Completa a la hora	Progresivos Rápidos: 12%
Días de Abstinencia: 5	Total Progresivos: 29%

En la Figura 1 se observan las tareas que se realizan durante el proceso de análisis y a los actores involucrados. A juicio de los expertos consultados, la realización de las pruebas para estimar la motilidad, viabilidad y morfología de los espermatozoides, el análisis para determinar la clasificación espermática y grados de severidad, la preparación del reporte y de las decisiones de recomendación y sugerencias, son tareas que requieren para su realización una gran carga cognitiva, objetividad y dominio del conocimiento del área porque son consideradas de alta complejidad. También la manipulación de las historias y la consulta de información del dominio (vía Internet), son de mucha utilidad porque ambas contribuyen a prestar un mejor servicio a profesionales y pacientes y permiten hacer un mejor uso de los recursos de información.

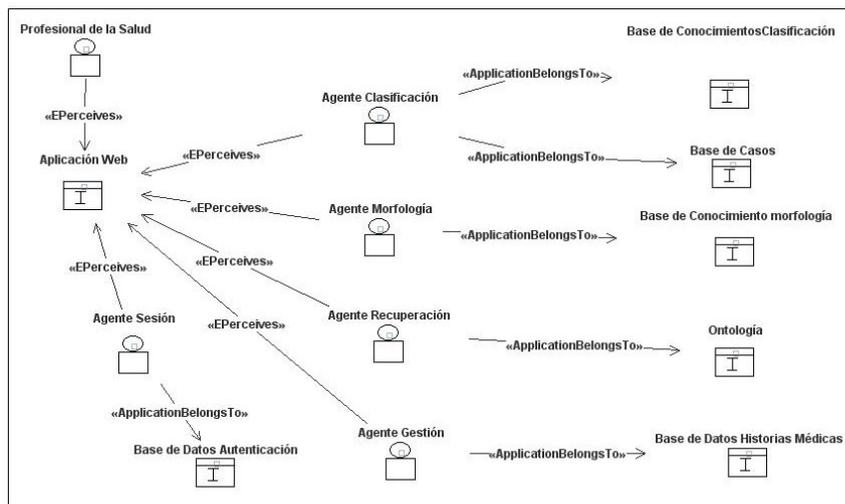


Figura 1. El proceso del análisis de líquido seminal humano

Se observa que para llevar a cabo el ALS, los profesionales deben articular una gran cantidad de elementos y conocimientos, razón por la cual alcanzar un diagnóstico adecuado y confiable se convierte en un proceso difícil. Un apoyo computacional que proporcione mecanismos para emular las tareas que requieren mayor carga cognitiva y experticia, facilidades de gestión e integración de información de los pacientes y que además permita observar el conocimiento del dominio, definitivamente sería de mucha utilidad para estos profesionales. Los SBA surgen como una opción de automatización para apoyar a los profesionales durante la realización del ALS.

ANÁLISIS Y DISEÑO DEL SISTEMA

Se diseñó un SBA denominado ANALISE (ANÁLisis de Líquido SEMinal), cuyo objetivo es emular algunas de las tareas que realizan los profesionales durante el proceso del ALS. El análisis se inició con la identificación de los

requerimientos de usuario, los cuales se corresponden con las tareas especificadas en la sección anterior. En la Tabla 3 se describen las funciones del sistema, el agente que la lleva a cabo y los recursos que utilizan.

Para completar la comunidad de agentes, se incorporó un Agente de Sesión, encargado de administrar las interacciones entre los usuarios y los agentes, y controlar y garantizar el acceso al sistema de manera segura. El recurso que utiliza este agente es una base de datos de autenticación.

Adicionalmente se desarrolló una aplicación Web mediante la cual los usuarios interactúan con el sistema (mecanismo a través del cual los agentes perciben y actúan).

El modelo de ambiente del sistema, que muestra a los agentes y los recursos que estos utilizan, se presenta en la Figura 2.

Tabla 3. Funciones, agentes y recursos de ANALISE

Función	Agente	Recursos
Clasificar la calidad de un eyaculado	Clasificación	-Base de casos de espermogramas -Base de conocimiento de clasificación (reglas para clasificar y recomendar)
Gestionar las historias médicas	Gestión	-Base de datos de historias médicas
Consultar y recuperar información del dominio	Recuperación	-Ontología (conocimiento estandarizado y organizado del dominio de la Andrología)
Analizar la morfología de las cabezas de espermatozoides	Morfología	-Base de conocimiento de morfología (imágenes de muestras de semen y vectores características)

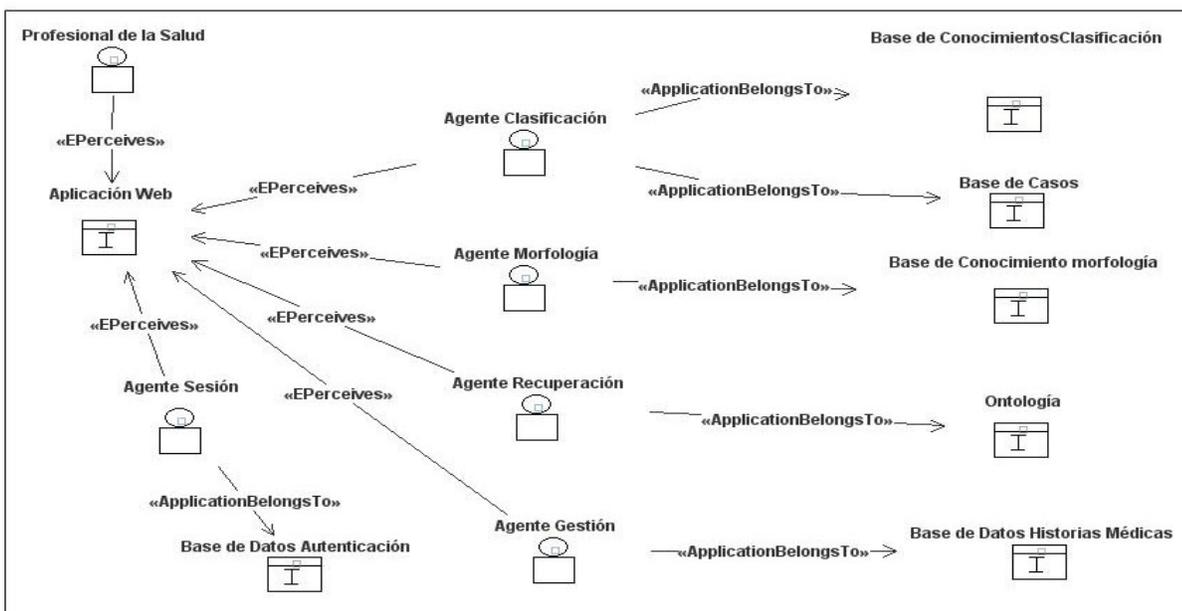


Figura 2. Modelo de ambiente de ANALISE

LOS AGENTES DEL SISTEMA ANALISE

Seguidamente se describen los agentes del sistema considerando objetivos, roles, tareas y recursos que utilizan.

Agente de sesión

Los usuarios interactúan con este agente a través de una aplicación Web que se ejecuta del lado del cliente. La administración de las interacciones, el control de acceso al sistema (a través de mecanismos de autenticación) y el establecimiento y mantenimiento de un área de trabajo, son responsabilidades del agente. El control de acceso que se implanta, garantiza la privacidad y seguridad de la información almacenada en las historias.

Agente de gestión

El agente de gestión (AG) es un agente de información que proporciona servicios de administración e integración de la información de los pacientes. Su objetivo es gestionar la base de datos de historias médicas. El AG permite crear y manipular las historias para obtener información de los pacientes tal como datos personales, espermogramas, diagnósticos, imágenes digitales, entre otros.

El AG suministra a los agentes del sistema que lo requieran, la información almacenada en las historias. Adicionalmente, calcula un conjunto de descriptores estadísticos (promedios, máximos y mínimos, desviación estándar, entre otros) de los parámetros de los espermogramas, a partir de los datos de los exámenes realizados en un centro de diagnóstico. Esta información constituye un valor agregado que permitirá analizar el comportamiento de una población particular.

Agente de recuperación y la ontología del dominio del ALS

El acceso a información médica actualizada, confiable y organizada, es una necesidad de los profesionales de la salud y de los pacientes. Por esta razón sería interesante que la gran cantidad de información del dominio de la Andrología, pudiese ser accedida fácilmente por esta comunidad. Pero, ¿cómo sería posible organizar la información de manera de hacerla accesible, confiable, precisa y de fácil entendimiento y representación? Definitivamente, las respuestas apuntan hacia las ontologías. A continuación se describe la Ontología del dominio del ALS diseñada para tal fin.

La ontología del ALS

Se diseñó y construyó una ontología del dominio del análisis del semen humano, que organiza, estructura y estandariza el conocimiento del dominio, el cual fue adquirido por los ingenieros de conocimiento mediante la revisión de documentación especializada, reuniones con expertos y participación en cursos y talleres dirigidos a profesionales de la Andrología. La ontología fue diseñada utilizando la metodología Methontology (Gómez-Pérez *et al.* 2004), y se formalizó con la herramienta Protégé (<http://protege.stanford.edu>).

La descripción amplia y detallada de los resultados obtenidos durante la realización de las actividades de especificación, conceptualización, formalización e implementación, sugeridas por la metodología, se muestra en Ramos *et al.* 2007.

En cada una de las actividades realizadas se verificaban y validaban los resultados obtenidos. Dado que esta ontología representa el primer desarrollo ontológico para el dominio de la Andrología, su evaluación se llevó a cabo aplicando el esquema propuesto en Ramos *et al.* 2009. La evaluación realizada permitió identificar errores e inconsistencias, así como omisiones en el vocabulario. La oportuna identificación e inmediata corrección de estos, permitió obtener un primer desarrollo ontológico de calidad.

El agente de recuperación

Este agente de información tiene como principal objetivo localizar y recuperar información en la ontología, para luego mostrarla de manera organizada a los usuarios a través de una interfaz Web especialmente diseñada para este fin (Figura 3).

El agente utiliza un algoritmo de comparación de cadenas de caracteres, que recupera información explícitamente codificada en la ontología (localizada en el nombre de una clase, instancia, propiedad o en las descripciones). Las búsquedas por comparaciones tienen como limitante, que alguna información que pueda resultar de interés para el usuario, tal vez no pueda ser recuperada ya que las relaciones entre los conceptos de la ontología no siempre se expresan de forma directa o explícita. Para solventar esta limitación, el agente dispone de un mecanismo de razonamiento basado en reglas, algunas de las cuales se describen en la siguiente sección, que le permite formular nuevas conclusiones a partir del conocimiento almacenado en la ontología.

Ontología del Análisis de Líquido Seminal Humano

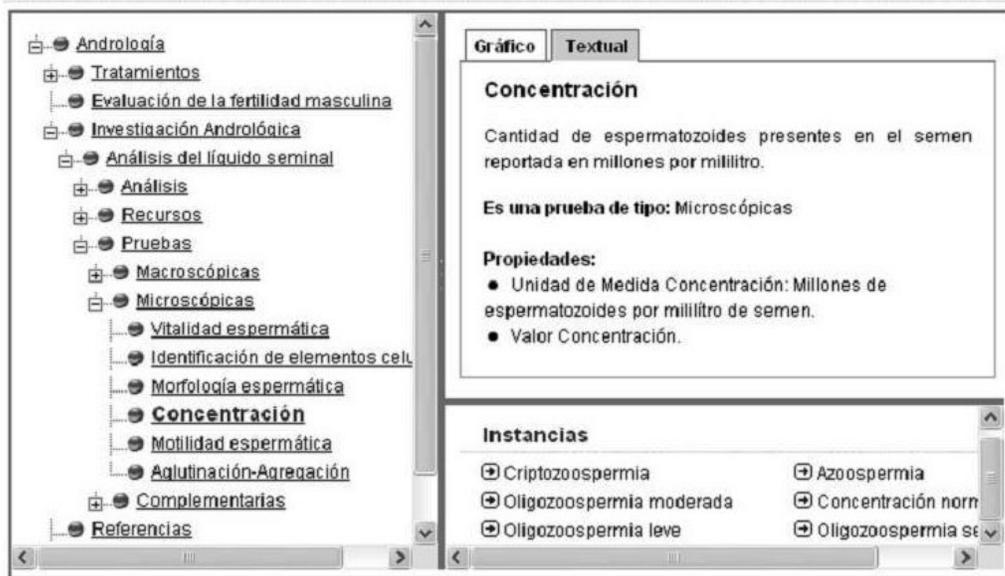


Figura 3. Recuperación de información desde la ontología del ALS

Agente de clasificación

Para clasificar la calidad de un eyaculado, el Agente de Clasificación (AC) actúa de manera similar a como lo haría un profesional; es decir, analiza las alteraciones que presenten las variables de un espermograma y determina la clase espermática y los grados de severidad. Además, emite sugerencias y comentarios de interés para el proceso de diagnóstico.

Cuando el profesional clasifica la calidad de un eyaculado, combina conocimiento fundamental del dominio, conocimiento empírico (casos resueltos) y conocimiento declarativo, el cual se expresa generalmente en forma de reglas. Para emular este proceso, el AC emplea un Mecanismo Híbrido Difuso de Razonamiento (MHDR) que combina estos tipos de conocimiento. El MHDR considera el uso de los paradigmas de razonamiento basado en reglas y razonamiento basado en casos y medidas de similitud difusa, porque estos se adaptan de manera adecuada al problema que se va a resolver; pues con las reglas es posible representar el conocimiento fundamental y declarativo, con los casos el conocimiento específico de los casos resueltos y con la lógica difusa los esquemas de pertenencia usados por los profesionales para clasificar los grados de severidad.

Para clasificar la calidad y determinar la severidad de una muestra, este agente deliberativo usa un Razonador Basado en Casos (RBC) y una base de casos de espermogramas, mientras que para sugerir y recomendar, utiliza un Razonador Basado en Reglas (RBR) y una base de conocimiento de reglas para sugerir. Estos mecanismos y recursos se describen a continuación:

a. Base de casos de espermogramas

La base de casos (BC) almacena espermogramas proporcionados por un laboratorio de fertilidad ubicado en Caracas-Venezuela. Luego del preprocesamiento de los espermogramas, el cual se realizó bajo la supervisión de los expertos, la BC quedó constituida por 2.663 casos, cada uno representado por un registro de dos componentes, tal como se muestra en la Tabla 4, donde el primero está relacionado con el resultado de las pruebas que serían las variables de entrada, y el segundo representa la clasificación de la muestra.

Tabla 4. Estructura de representación de casos de la BC

Resultado Pruebas	Identificador, Volumen, Concentración, % Inmóviles Vivos, Formas Normales, Aspecto, pH, Color, Viscosidad, Aglutinación, Progresivos_Rápidos (PR), Progresivos_Moderados (PM), No Progresivos, Inmóviles Vivos, Inmóviles Muertos, Formas Anormales, Cabezas Grandes, Cabezas Pequeñas, Cabezas Redondas, Cabezas Duplicadas, Amorfos, Defecto Cuello, Defecto Cola,
Solución	Clasificación Espermática, Grado de Severidad, Observaciones, Sugerencias

Para determinar el índice que va a ser utilizado para recuperar los casos, se usaron diferentes técnicas de selección de atributos, para identificar las variables más importantes que influyen o determinan la solución. Específicamente, se aplicaron técnicas filtro, utilizando criterios de evaluación como ganancia de información y chi cuadrado, entre otras (Piramuthu, 2004), y técnicas por modelos (wrapper) entrenando una máquina de soporte vectorial, con los parámetros utilizados por defecto por la herramienta WEKA 3.4.1.1 (<http://www.cs.waikato.ac.nz/ml/weka/>). Los resultados fueron verificados y validados por los expertos, quedando el índice constituido por las siguientes variables: Volumen, Concentración, PR+PM, %Formas Normales y % Inmóviles Vivos.

b. Base de conocimiento de clasificación

Esta base está constituida por dos grupos de reglas. El primero considera las reglas utilizadas para recuperar casos, las cuales fueron derivadas aplicando el algoritmo de árboles de decisión C4.5 (Larose, 2005) a la BC, para un total de nueve reglas que permiten identificar las cincuenta y un clases espermáticas posibles. El segundo grupo de reglas, usadas para sugerir y recomendar, toman en cuenta un conjunto de variables que no determinan la clasificación, pero que aportan conocimiento para el diagnóstico. Ambos grupos de reglas fueron verificadas y validadas por los expertos.

En la Tabla 5 se muestran algunas de las reglas usadas por el agente para recuperar casos.

Tabla 5. Reglas de recuperación de casos

Regla 1	Si Volumen < 2 ml Entonces Clasificación = “Hiperespermia”
Regla 2	Si Cantidad de Espermatozoides < 20 millones Entonces Clasificación = “Oligozoospermia”
...	...
Regla 9	Si Espermatozoide Formas Normales < 15% Entonces Clasificación = Teratozoospermia”

Una muestra de las reglas usadas para sugerir y recomendar, se presentan en la Tabla 6.

Tabla 6. Reglas para sugerir y recomendar

Regla 1	Si Tiempo de Recolección > 60 minutos Entonces Sugerencia = “Ya que el tiempo transcurrido desde que se colectó la muestra ha superado los 60 minutos, no se debe realizar licuefacción ni estimar la viscosidad.”
...	...
Regla 16	Si Células Redondas > 6 Entonces Sugerencia = “Debido a la presencia de células redondas en la muestra, es conveniente realizar un test de peroxidasa”

c. El mecanismo híbrido difuso de razonamiento

El razonador basado en casos utiliza un esquema de recuperación en dos etapas: en la primera, empleando las reglas de recuperación, recupera los casos cuya clasificación espermática coincida con la clase del caso de entrada, obteniéndose una base de casos reducida que contiene sólo los casos que pertenecen a dicha clase; en la segunda etapa, utilizando el algoritmo de K-vecinos más cercanos (*K Nearest Neighbors* - KNN) y una medida de similitud difusa (Hernández *et al.* 2004), se recuperan de la base de casos reducida los casos cuyas sub-clasificaciones (grados de severidad) sean más parecidos al caso de entrada. La medida de similitud difusa utilizada se explica seguidamente:

Sea X_{nuevo} un nuevo caso a resolver y Y_{bc} un caso de la base de casos reducida, donde:

$$X_{nuevo} = (x_1, x_2, x_3, x_4, x_5) \quad \text{y} \quad Y_{bc} = (y_1, y_2, y_3, y_4, y_5)$$

Los $(x_1, x_2, x_3, x_4, x_5)$ y $(y_1, y_2, y_3, y_4, y_5)$ son los valores de las variables del índice que caracterizan ambos casos.

La similitud entre ellos se calcula como:

$$SIM(X_{nuevo}, Y_{bc}) = \sum d_i \tag{1}$$

donde: $d_i = \max[\min(\mu_{severa}(x_i), \mu_{severa}(y_i)),$

$$\min(\mu_{moderada}(x_i), \mu_{moderada}(y_i)),$$

$$\min(\mu_{leve}(x_i), \mu_{leve}(y_i))]$$

Los conjuntos difusos utilizados, se determinaron a partir de los datos calculando, para cada variable del índice con

grado de severidad, el valor mínimo y el valor máximo, así como el punto medio de los datos asociados con cada una de las etiquetas lingüísticas (severa, moderada y leve). Para medir el rendimiento del MHDR, se realizaron un conjunto de pruebas aplicando dos variaciones del algoritmo de KNN, por mayoría y basado en la distancia promedio, utilizando la medida difusa descrita con la ecuación (1), con funciones de pertenencia triangular y trapezoidal.

Para determinar la influencia del grado de solapamiento entre conjuntos difusos adyacentes, éste se varió en una proporción de -10%, -5%, 5% y 10%. La Tabla 7 muestra los resultados obtenidos durante el entrenamiento, utilizando validación cruzada (Hernández *et al.* 2004) de 10 particiones, con una vecindad fija de tamaño 3. Esta vecindad fue obtenida experimentalmente y fue la que arrojó mejores resultados.

La estructura de dos niveles usada por el MHDR, produce una parte de la solución formada por la clasificación esperámica de la muestra y los grados de severidad, la cual complementa el Razonador Basado en Reglas suministrando las sugerencias y observaciones para el caso.

Tabla 7. Resultados de las pruebas sobre la BC con medida de similitud difusa

Medida	Promedio porcentaje de aciertos	
	Mayoría	Distancia promedio
Triangular -5%	76.42	86.87
Triangular -10%	76.50	87.54
Triangular 5%	74.72	86.77
Triangular 10%	74.43	85.63
Trapezoidal -5%	77.75	90.57
Trapezoidal -10%	77.80	90.79
Trapezoidal 5%	77.83	91.26
Trapezoidal 10%	78.03	91.33

La arquitectura del MHDR, esquematizada según las fases del ciclo del RBC (recuperación, adaptación, revisión y retención) se observa en la Figura 4.

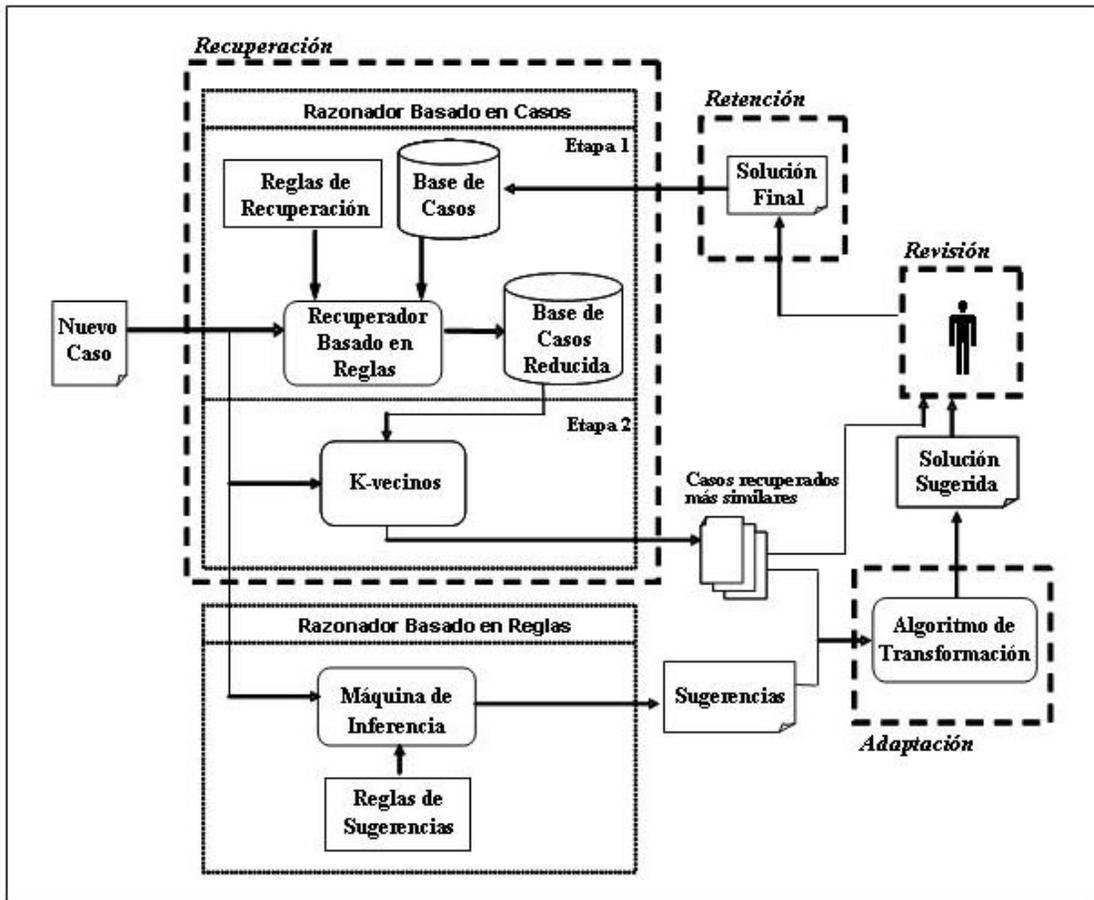


Figura 4. Arquitectura del mecanismo híbrido difuso de razonamiento

El rendimiento alcanzado por el MHDR (91,33% de aciertos), evidencia su capacidad para clasificar la calidad de una muestra. Estos resultados proporcionan a los profesionales apoyo oportuno durante el diagnóstico de la calidad de un eyaculado (Ramos *et al.* 2008a).

Agente de morfología

Analizar la morfología espermática es considerada una de las tareas más complicadas y costosas del ALS. El Agente de Morfología (AM) es el encargado de emular esta tarea (restringida a la morfología de las cabezas de los espermatozoides) utilizando técnicas de procesamiento de imágenes y paradigmas de aprendizaje de máquinas como las Redes Neuronales (RN) y las Máquinas de Soporte Vectorial (SVM), (Hernández *et al.* 2004), a partir de imágenes digitalizadas de muestras de semen.

Las imágenes se capturaron a partir de quince láminas de muestras de semen, para un total de 160 imágenes que contenían cerca de 1600 espermatozoides. Cada espermatozoide fue segmentado y caracterizado, utilizando para esto último un conjunto de algoritmos diseñados específicamente para extraer características de las cabezas, similares a las sugeridas por la OMS para evaluar la morfología espermática (forma, largo y ancho de la cabeza, área total de la cabeza y área ocupada por el acrosoma). Una amplia descripción de estos algoritmos se encuentra en Ramos *et al.* 2008b).

Los vectores de características extraídos fueron etiquetados por los expertos en dos clases: Espermatozoide Normal y Espermatozoide Anormal. Estos vectores fueron usados para estimar diferentes modelos de clasificación utilizando RN y SVM, con la finalidad de seleccionar aquel que exhibiera mejor rendimiento y capacidad de respuesta. En el caso de las RN se utilizó una red Perceptron Multicapas entrenada con backpropagation, con quince neuronas en la capa de entrada y una capa oculta con diez neuronas (este último parámetro fue obtenido experimentalmente); las pruebas fueron realizadas con la ayuda de la herramienta Weka 3.4.1.1 Para el caso de las SVM se emplearon funciones núcleos RBF y para el entrenamiento se usó el software “OSU Support Vector Machines Toolbox” versión 3.0 (Ma & Zhao, 2002), en el ambiente de programación Matlab 6.5.

El archivo de datos utilizado para el entrenamiento, estaba constituido por 960 registros (vectores de características de los espermatozoides), 626 estaban clasificados como anormales y 334 como normales. Para determinar el rendimiento de clasificación para ambos paradigmas, se realizaron 5 validaciones cruzadas de 10 particiones y

se tomaron los valores promedios sobre el porcentaje de aciertos. Estos resultados se muestran en la Tabla. 8.

Tabla 8. Resultados del entrenamiento con la red neuronal y con la SVM

Paradigma	Porcentaje de aciertos
Perceptron Multicapas	76,74
SVM	76,04

Es importante destacar que la mayor cantidad de errores se presenta al clasificar como anormales, registros que fueron etiquetados como normales. Esto puede deberse a que el número de espermatozoides normales en la base de datos, es relativamente pequeño dado el origen de las muestra, porque ellas pertenecen a pacientes que acuden a centros especializados, en general, por problemas de fertilización, aspecto que hace presumir la existencia de defectos de morfología.

Aun cuando el rendimiento obtenido en la clasificación no fue satisfactorio, es evidente que esta primera aproximación muestra el potencial de combinar técnicas de procesamiento de imágenes y paradigmas de aprendizaje de máquinas para analizar la morfología de la cabeza de los espermatozoides a partir de imágenes. Es por esta razón que este agente continúa en desarrollo y actualmente se ensayan otros mecanismos para segmentar los espermatozoides, se mejoran los algoritmos de extracción de características, se evalúan otras posibilidades de descripción morfológica y se están adquiriendo imágenes con mayor resolución, entre otras acciones. Además, se están evaluando otras técnicas de aprendizaje que tomen en cuenta conjuntos de datos con clases no balanceadas.

La Figura 5 muestra la arquitectura del sistema ANALISE, la manera como los agentes interactúan y colaboran entre sí para alcanzar los objetivos para los cuales fueron diseñados.

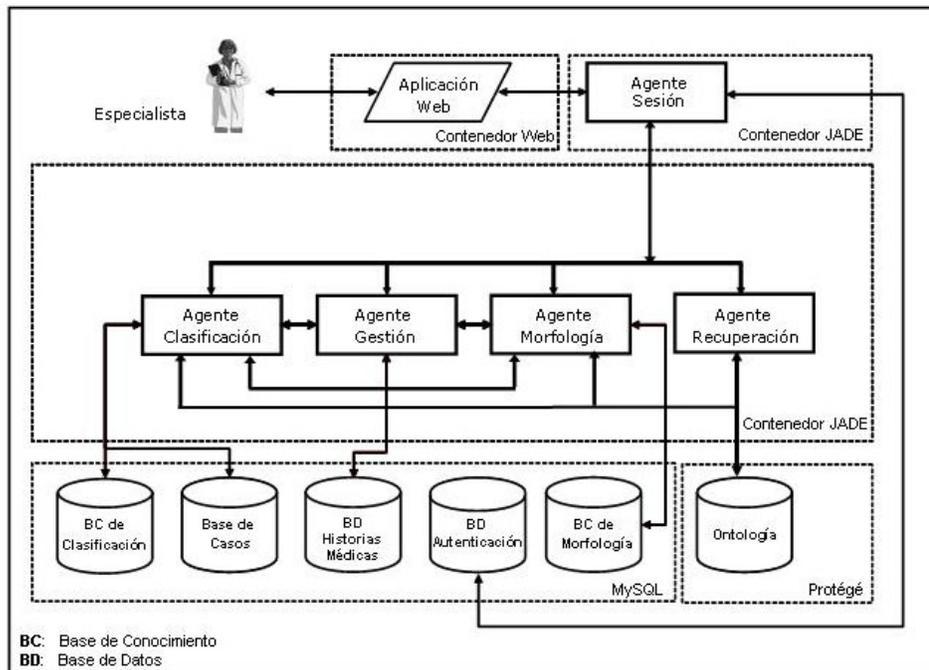


Figura 5. Arquitectura del sistema ANALISE

CONCLUSIONES

La tecnología de agentes, representada en esta investigación por la estructura organizacional de los agentes (roles, tareas, entre otros), las técnicas de razonamiento y aprendizaje, las metodologías y herramientas de desarrollo, los esquemas de representación y estructuración del conocimiento (ontologías), facilitaron el diseño y desarrollo de una solución computacional para apoyar a los profesionales de la salud durante la realización del análisis del semen humano.

El proceso de analizar el líquido seminal fue adecuadamente dividido en subproblemas para que fuesen resueltos por los agentes del sistema ANALISE.

Los resultados alcanzados fueron satisfactorios porque el sistema fue capaz de: a) clasificar la calidad de muestras de semen (clasificación espermática y grados de severidad), b) gestionar la información almacenada en las historias de los pacientes, c) proveer información organizada y estandarizada del conocimiento del dominio de la Andrología y d) en una primera aproximación, caracterizar las cabezas de los espermatozoides a partir de imágenes digitalizadas, usando algoritmos específicamente diseñados para este fin, los cuales permitieron realizar el análisis morfológico.

La arquitectura propuesta podría ser usada para el desarrollo de una plataforma para el análisis de otros líquidos biológicos cuyo análisis se lleve a cabo de manera similar al líquido seminal.

REFERENCIAS

- BAJO, J., PAZ, J., PAZ, Y., CORCHADO, J. (2009). *Integrating case-based planning and RPTW neural networks to construct an intelligent environment for health care*. Expert Systems with Applications. Volume 36, Issue 3, Part 2. 5844-5858.
- CORCHADO, J., BAJO, J., PAZ, Y., TAPIA, D. (2008). *Intelligent environment for monitoring Alzheimer patients, agent technology for health care*. Decision Support Systems Volume 44. Issue 2. January. 382-396.
- GANDON, F., POGGI, A., RIMASSA, G., TURCI, P. (2002). *Multi-Agent Corporate Memory Management System. In Engineering Agent Systems: Best of "From Agent Theory to Agent Implementation (AT2AI)-3"*. Journal of Applied Artificial Intelligence, Volume 16, Number 9-10/October Taylor & Francis.
- GÓMEZ-PÉREZ, A., FERNÁNDEZ-LÓPEZ, M., CORCHO, M. (2004). *Ontological Engineering*. London: Springer-Verlag.
- GUPTA, S. & PUJARI, S. (2009). *A multi-agent system (MAS) based scheme for health care and medical diagnosis system*. International Conference on Intelligent Agent & Multi-Agent Systems. IAMA 2009. Print ISBN: 978-1-4244-4710-7. Chennai, Indian.
- HEINE, C., HERRLER, R., STEFAN, K. (2004). *Agent-Based*

- Optimization and Management of Clinical Processes.* Proceedings of the 16th European Conference on Artificial Intelligence (ECAI'04). The 2nd Workshop on Agents Applied in Health Care.
- HENDERSON-SELLER, B. & GIORGINI, P. (2005). *Agent-Oriented Methodologies: An Introduction.* En Henderson-Seller, B. Giorgini, P. (Eds). Agent-Oriented Methodologies. (Chapter I). IDEA Group Publishing. ISBN: 1-5914-586-6
- HERNÁNDEZ, J., RAMÍREZ, MA J., FERRI, C. (2004). *Introducción a la Minería de Datos.* ISBN: 8420540919. Prentice-Hall.
- HUANG, M., CHEN, M., LEE, S. (2007). *Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis.* Expert Systems With Applications. Vol. 32. 856-867.
- ISERN, D., SÁNCHEZ, D., MORENO, A. (2007). *Agents and clinical guidelines: Filling the semantic gap.* En Angulo, C. y Godó, L. (Eds.). Artificial Intelligence Research and Development (pp. 67-76), CCIA 2007. IOS Press. Frontiers in Artificial Intelligence and Applications. Vol. 126. Andorra la Vella, Andorra.
- JENNINGS, N. & WOOLDRIDGE M. (2000). *Agent-Oriented Software Engineering.* Artificial Intelligence. Vol 117. 277-296.
- KUMAR, M. & SYAMALA, M. (2007). *A Multi-agent Medical System for Indian Rural Infant and Child Care.* Twentieth International Joint Conference on Artificial Intelligence. Hyderabad, India.
- LAROSE, D. (2005). *Discovering Knowledge in Data. An Introduction to Data Mining.* Wiley Interscience.
- LUCK, M., MCBURNEY, P., SHEHORY, O., WILLMOTT, S. (2005). (Ed): *Agent Technology: Computing as Interaction. A Roadmap for Agent Based Computing.* The AgentLink Community, University of Southampton. ISBN 085432 845 9.
- MA, J. & ZHAO, Y. (2002). *OSU Support Vector Machines Toolbox, version 3.0.* Recuperado en Marzo de 2008 en: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- MORTIMER, D. (1994). *Practical Laboratory Andrology.* Oxford University Press.
- OMS- ORGANIZACIÓN MUNDIAL DE LA SALUD. (2001). *Manual de Laboratorio para el examen del semen humano y de la interacción entre el semen y el moco cervical.* Madrid, España: Editorial Médica Panamericana S.A. Abril. Cuarta Edición.
- PAVÓN, J., GÓMEZ-SANZ, J. Y FUENTES, R. (2005). *The INGENIAS Methodology and Tools.* En Henderson-Seller, B. y Giorgini, P. (Eds). Agent-Oriented Methodologies. (Chapter IX). IDEA Group Publishing. ISBN: 1-5914-586-6.
- PIRAMUTHU, S. (2004). *Evaluating feature selection methods for learning in data mining applications.* European Journal of Operational Research. 156. 483-494
- RAMOS E., NÚÑEZ, H., CASAÑAS, R. (2008a). *A Fuzzy Hybrid Intelligent System for Human Semen Analysis.* En Hector Geffner, Rui Prada, Isabel Machado Alexandre, Nuno David (Eds.): Advances in Artificial Intelligence - IBERAMIA 2008, 11th Ibero-American Conference on AI, Lisbon, Portugal, 2008. Proceedings Lecture Notes in Computer Science 5290 Springer 2008, ISBN 978-3-540-88308-1.
- RAMOS, E., NÚÑEZ, H., CASAÑAS, R., CALDERÓN, E., FERNÁNDEZ, V. (2008b). *Neural Networks and Support Vector Machines for the Classification of the Morphology of Human Sperm Heads from Images.* Jornadas Chilenas de Computación. Workshop de Inteligencia Artificial. Chile.
- RAMOS, E., NÚÑEZ, H., PEREIRA, Y., CASTRO, M., CASAÑAS, R. (2007). *Aplicación de Visualización de una Ontología para el dominio del Análisis del Semen Humano.* Ingeniería y Ciencia, Vol. 3, número 5. Junio. Medellín. Colombia. ISSN 1794-9165.
- RAMOS, E., NÚÑEZ, H., CASAÑAS, R. (2009). *Un Esquema para Evaluar Ontologías Únicas para un Dominio de Conocimiento.* Enl@ce: Revista Venezolana de Información, Tecnología y Conocimiento. Número 1. Año: 2009. Enero-Abril. ISSN: 1690-7515. pp 57-72.
- RUSSELL, S. & NORVIG, P. (2004). *Inteligencia artificial: Un enfoque moderno.* Segunda Edición. PEARSON Prentice Hall. ISBN. 84-205-4003-X. España.
- SU, CH. & WU, CH. (2009). *JADE implemented mobile multi-agent based, distributed information platform for pervasive health care monitoring.* Applied Soft Computing, In Press, Corrected Proof, Available online 22 November 2009.