

PRESENTACIÓN Y LÍNEAS DE INVESTIGACIÓN INSTITUCIONALES



Instituto de Investigación en Tecnología Informática Avanzada (INTIA)

Expositor: Mg. TOSINI, Marcelo

<http://www.intia.exa.unicen.edu.ar>

marcelo.tosini@gmail.com

INTRODUCCIÓN

El INTIA tiene su sede en la Facultad de Ciencias Exactas de la Unicen, en la ciudad de Tandil. Alberga a 28 investigadores de los cuales 22 son docentes con cargos profesoriales y 23 tienen sus posgrados. Cuenta además con estudiantes de posgrado realizando actualmente becas Conicet y CIN.

Las investigaciones de la institución se agrupan en cuatro grandes líneas: Bases de Datos y Procesamiento de Señales, Sistemas Digitales y Tecnología de Software, e Informática de Gestión; las cuales serán descritas a continuación.

BASES DE DATOS Y PROCESAMIENTO DE SEÑALES

Modelado de datos no convencionales. En el modelado de datos el Diagrama de Entidades y Relaciones Extendido es una pieza clave, a partir del cual, mediante procesos de transformaciones se genera un esquema físico de datos, basado, en su gran mayoría, en un modelo Objeto-Relacional. En dicho proceso existe una variedad de problemas, originados por la falta de herramientas y metodologías que permitan expresar la semántica de determinadas situaciones.

Para el caso de datos morfométricos se plantean nuevas perspectivas y necesidades de procesamiento de datos, que difieren de aquellas ya firmemente establecidas en los sistemas que procesan datos convencionales; más específicamente, la morfometría geométrica utiliza coordenadas cartesianas en 2D o 3D de puntos anatómicos específicos que pueden ser claramente identificados y registrados en todos los organismos cuya forma se quiere analizar y/o comparar.

Técnicas Cuantitativas Orientadas al Reuso Semántico de Modelos de Requisitos.

Los modelos LEL, Escenarios Actuales, Escenarios Futuros y Especificación de Requisitos (SRS) son construidos para dar soporte al proceso de definición de los requisitos del software. Sin embargo, su mera observación objetiva muestra que contienen más información que la que explícitamente se indica. Por ejemplo, la existencia de dos símbolos del LEL que son claramente sinónimos parciales, muestra que existe una jerarquía cuyos miembros no pertenecen al LEL en su totalidad. Esto indica que el Universo de Discurso ha fallado en plasmar en su vocabulario todas las abstracciones posibles que emergen de su problemática, o que la captura del glosario ha sido incompleta. Para poder extraer esta y muchas

otras informaciones de los modelos del proceso de Ingeniería de Requisitos, se requiere construir grafos, contar referencias, calificar relaciones y otros mecanismos que aún se desconocen. Este es el eje de las actividades de esta línea de investigación, ya que se pretende sistematizar el metaconocimiento disponible y enfatizar el uso de diferentes técnicas basadas en estrategias cuantitativas (figura 1).

Clusterización en imágenes multiespectrales. Obtener información de imágenes satelitales requiere ordenar y agrupar una gran cantidad de datos multidimensionales. Las técnicas de *Clustering* implementadas en computadoras hasta la fecha, brindan soluciones aceptables; sin embargo, tienden a ser lentos y pueden requerir intervención humana en muchos pasos del proceso. Además, algunas de estas técnicas dependen en gran medida de las inicializaciones específicas y contienen pasos no deterministas, lo que lleva a un largo proceso de prueba y error hasta que se encuentre un resultado adecuado. El grupo ha desarrollado a *FAUM: Fast Autonomous Unsupervised Multidimensional*, un algoritmo de *Clustering* automático que puede descubrir agrupaciones naturales en *big data*. Se basa en generar diferentes histogramas multidimensionales, definidos como *Hyper*-histogramas, y elegir uno para obtener la información. *FAUM* tiene como objetivo optimizar los recursos proporcionados por una computadora moderna. Cuando se inició la presente investigación, la idea era desarrollar un método de inicialización determinista de *K-Means*, sin proporcionar el número de clústeres para encontrar y aplicable a las Imágenes Satelitales.

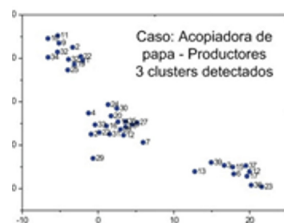


Figura 1. Agrupamientos en el LEL

Durante la fase de prueba, se encontró que el algoritmo también se puede extender a un algoritmo de *Clustering* cerrado en sí mismo. Además, se puede generalizar fácilmente para procesar cualquier gran conjunto de datos multidimensionales, con un rendimiento notable en comparación con *K-Means*. Aunque el algoritmo es autónomo, puede ajustarse manualmente, si se desea. *FAUM* trata el proceso de *Clustering* como una sucesión de pasos. Cada paso extrae información relevante del conjunto de datos de entrada, generando un conjunto de datos nuevo y más pequeño, que se utilizará en el siguiente paso. A la fecha se han publicado los primeros dos pasos de *Clustering* de *FAUM*, que obtiene resultados aceptables con conjuntos de datos que contienen grupos simétricos esféricos disjuntos, similares a los *K-Means*. En este punto, *FAUM* se puede considerar un método de agrupamiento lineal. Se están investigando más pasos y se presentarán en el futuro. Estos pasos podrían permitir a *FAUM* abordar clústeres no lineales, o incluso incluir otras técnicas no lineales actualmente en uso, como *COLL*, *MEAP*, entre otras, que podrían aprovechar la reducción del tamaño del conjunto de datos.

Sistemas de asistencia al diagnóstico y al tratamiento. La determinación de la dosis de radiación absorbida en tejido humano es de suma importancia para lograr un tratamiento de radioterapia eficaz. El método más preciso para estimar la dosis es el cálculo basado en Montecarlo pero los programas que realizan este cálculo han debido optar por alguna de las variantes en el compromiso, aún no resuelto apropiadamente, entre la calidad de la estimación y el costo temporal del cálculo. Muchas de las técnicas existentes de reducción de tiempo se basan en una simplificación del problema y acarrear una pérdida de calidad en los resultados. Se han aplicado técnicas de cálculo directo, el precálculo y la paralelización, que permiten reducir el tiempo de cálculo sin perder calidad en los resultados, realizando un cálculo completo sin simplificaciones.

La digitalización de las diferentes modalidades de imágenes médicas y su disponibilidad facilita el desarrollo de herramientas de procesamiento digital de imágenes con el fin de brindar soluciones de asistencia al diagnóstico. En este sentido se han aplicado redes neuronales convolucionales (*CNN*) en una variedad de problemas de segmentación y clasificación en imágenes médicas. Estas imágenes presentan varios desafíos para la implementación de *Deep Learning*, tales como la existencia de diferentes modalidades, la dificultad en el etiquetado por expertos (*Ground Truth*) y los defectos propios de las imágenes médicas. Se ha trabajado en la implementación de técnicas basadas en Deep Learning principalmente en problemas de segmentación de la pared vascular en imágenes de ultrasonido intravascular, en caracterización de hueso trabecular, en conteo de células en imágenes hiperespectrales de microscopía de fluorescencia y otros problemas. Además, se está trabajando en el reconocimiento de patrones de uso en herramientas de rehabilitación cognitiva. Para esto se están desarrollando un conjunto de herramientas y se aplican técnicas de procesamiento de imágenes para la detección de patrones, por ejemplo de detección de la mirada.

SISTEMAS DIGITALES

El grupo se centra en la investigación y el desarrollo de soluciones embebidas, desarrollando sobre diversas plataformas, como microcontroladores y *FPGAs*. El laboratorio tiene como objetivo la formación continua de recursos humanos en la temática, manteniendo un nivel tecnológico competitivo y la participación, junto a empresas privadas y públicas, en la presentación de proyectos, la asesoría técnica y la prestación de servicios específicos.

El Laboratorio surge a partir de las intenciones comunes de integrantes del Instituto en desarrollar soluciones de sistemas embebidos a problemas particulares, así como también realizar investigaciones, ya sea de optimización o aplicación de la tecnología actual. El interés regional en encontrar soluciones a medida, a costos competitivos potencia la creación de este laboratorio. Sus integrantes poseen más de 15 años de experiencia en la temática y sus experiencias particulares se complementan para ofrecer servicios de investigación y desarrollo sobre diferentes plataformas. Además, han participado en diversos proyectos de empresas particulares y organismos públicos, así como también, han realizado actividades académicas en el área, con una activa participación en congresos nacionales relacionados.

El laboratorio centra sus investigaciones en: diseños basados en sistemas microprogramados, diseños basados en lógica programable, diseños basados en *SoC FPGA*, desarrollo de *cores IP*, verificación funcional, aceleración y optimización de Sistemas Embebidos, análisis y diseño de algoritmos complejos para Sistemas Embebidos, desarrollo de soluciones *IoT*, entre otros.

TECNOLOGÍAS DE SOFTWARE

Gran parte de los sistemas de información vitales en organizaciones de nuestro medio fueron implementados hace varios años con tecnologías que no están alineadas con los actuales objetivos estratégicos de las organizaciones. Estos sistemas, conocidos como *legacy* (heredados), involucran software, hardware, procesos de negocio y estrategias organizacionales. En general no están documentados o, si lo están, sus especificaciones no reflejan los cambios de requerimientos que se dieron a través de los años y solo el código engloba la historia de su evolución. Además, la entrada en escena de nuevas tecnologías, por ejemplo móviles, motiva la creciente demanda de modernización de sistemas desarrollados más recientemente

por ejemplo, con tecnologías orientadas a objetos. Nuevos enfoques de desarrollo de software enmarcados en lo que se referencia como *MDD (Model Driven Development)* podrían dar respuesta a esta demanda, sin embargo, no están lo suficientemente difundidos en nuestro medio que en general opta por nuevos desarrollos que conllevan un alto riesgo y no son rentables. Se propone en esta investigación una integración de *ADM* con técnicas clásicas de ingeniería inversa de análisis estático y dinámico, con técnicas rigurosas de metamodelado para controlar la evolución de software hacia nuevas o actuales tecnologías.



Figura 2. Esquemas de modernización de software partiendo de metodologías ADM hasta esquemas actuales de metamodelado

La idea es potenciar la productividad mediante la definición de las bases para el desarrollo de herramientas que aumenten el grado de automatización en procesos de modernización de software. Se analizan diversos escenarios de modernización, poniendo el énfasis en la modernización de software desarrollado en C/C++ y Java a fin de adaptarlo a tecnologías móviles.

El desarrollo de software alineado a estos nuevos paradigmas requiere adaptar software *legacy*. Los desafíos en este tipo de migraciones tienen que ver con la proliferación de plataformas móviles de desarrollo y con la carencia de procesos reutilizables y sistemáticos con un alto grado de automatización que reduzcan riesgos, tiempo y costos. Con respecto al primer desafío, la situación ideal es trabajar con desarrollos multiplataforma que permitan implementar una aplicación y desplegarla automáticamente sobre diversas plataformas. Por otra parte, *MDE (Model Driven Engineering)*, una disciplina de desarrollo de software que enfatiza en el uso de modelos en diferentes niveles de abstracción, puede aportar a la industrialización de la modernización de software. Se propone en esta investigación combinar *MDE* y desarrollos multiplataforma a fin de adaptar software *legacy* a nuevas tecnologías, como *Mobile Computing*, *IoT* y *Cloud Computing*.

GESTIÓN

Análisis y elaboración de datos para sistemas de indicadores de ayuda social.

El objetivo del proyecto es desarrollar un sistema de indicadores para gestionar, el sistema de ayuda solidaria KOINONIA, a partir de la aplicación de algoritmos y técnicas de Inteligencia Artificial y de Gestión del Conocimiento sobre los datos disponibles y detectables en la red de asistencia solidaria y en la población asistida, mediante tecnologías WEB.

Herramientas para la toma de decisiones aplicado a la mejora de indicadores académicos. Programa PEFI. Tableros de control e integrador de fuentes de datos.

El objetivo del proyecto presentado es el diseño y desarrollo de procedimientos y herramientas de software que asistan a la mejora de indicadores académicos en las facultades de ingeniería de la Universidad Nacional del Centro de la Provincia de Bs. As. en cuanto a los Programas de Ingreso, Permanencia y Graduación.