

Directorio

Sociedad Mexicana de Ciencias de la Computación

Presidente

Alfredo Sánchez

Vicepresidente

Carlos Coello

Secretario

Manuel Romero

Tesorero

Antonio Sánchez

Vocal

Amalia Duch

Vocal

Eduardo Morales

Vocal

Antonio Martínez

Vocal

Carlos Avilés

Vocal

Marcelo Mejía

boletín de la SMCC

Consejo Editorial

Miembros de la Mesa Directiva

Pablo Noriega B. V.

Atocha Aliseda

Editora

Amalia Duch

Diseño y edición dtp

pixelarte

Publicación estacional de la Sociedad Mexicana de Ciencias de la Computación. Casa No. 24 UDLA-P. Cholula, Pue. C.P. 72820. Tels.: (52)(2) 2-29-20-02. URL: www.lania.mx/~smcc e-mail: smcc@xalapa.lania.mx

Certificado de Reserva al título en Derechos de Autor: en trámite.

Editorial

Queridos Lectores:

Tienen en sus manos (o en sus monitores) el primer número del boletín de la Sociedad Mexicana de Ciencias de la Computación (SMCC). Esperamos que lo disfruten. Está hecho para ustedes con entusiasmo y con el convencimiento de que es una publicación útil no sólo para los miembros de la SMCC, sino para la comunidad informática en general.

Estamos convencidos de que este boletín es útil, pues pretendemos que a través suyo mejore la comunicación y la interacción entre los distintos miembros de la comunidad de ciencias de la computación. Queremos que se consolide como un medio de libre expresión para investigadores y estudiantes, quienes pueden estar seguros de que aquí tienen un espacio para difundir, sugerir, aclarar u opinar sobre los diversos temas de interés de nuestra comunidad y, desde luego, para definir esos temas y sacarlos a la luz pública.

Queremos que las distintas instituciones de educación superior en computación o áreas afines, los institutos de investigación, las sociedades o asociaciones afines a la SMCC y empresas del ramo, tengan aquí un espacio para darse a conocer, expresar sus distintos proyectos y actividades, entrar en contacto con otras instituciones relacionadas, con investigadores y con alumnos.

Seguramente todo ello será posible gracias a la participación y colaboración de nuestros asociados y de los miembros de la comunidad en general. Esperamos que cada tres meses (o antes, si la participación de ustedes es abrumadora y las finzas así lo permiten) puedan tener en sus manos un nuevo número de este boletín.

Agradecemos a todos aquellos que participaron en la elaboración de este primer número. Sabemos que podemos y debemos mejorarlo edición tras edición, y para ello todas las sugerencias de nuestros lectores son bienvenidas.

Amalia Duch

ENC '99

pp. 2-6

José Cuenca

QEPD

p. 7

Entrevista con
Vladimir Estivill-Castro

p. 9

Hallando los temas principales en un
artículo en español/ Adolfo Guzmán

p. 14

El II Encuentro Nacional de Computación ENC'99, se celebró los días 12, 13, 14 y 15 de septiembre de 1999 en la ciudad de Pachuca, Hidalgo, teniendo como sede la Universidad Autónoma del Estado de Hidalgo.

El objetivo fundamental del ENC'99 fue el de promover y difundir la investigación y desarrollo del área de las ciencias computacionales en nuestro país.

El primer encuentro, ENC'97, tuvo lugar en la ciudad de Querétaro, Querétaro, con la asistencia de más de 300 personas.

Al ENC'99 asistieron alrededor de 300 personas.

El ENC'99 estuvo organizado alrededor de siete talleres por área de especialidad y por una serie de cursos tutoriales y conferencias magistrales alrededor de los distintos ámbitos de competencia. Los talleres presentados en esta ocasión fueron los siguientes: Aprendizaje, búsqueda, clasificación y optimización con enfoques heurísticos • Bases de Datos • Ingeniería de Software • Interacción mediada por computadora • Lógica y Computación • Sistemas Distribuidos y Paralelos • Graficación, Visión y Robótica.

Para mantenerte al tanto, en este número de nuestro boletín los coordinadores de cada uno de los talleres mencionados presentan una reseña de lo que fue el ENC'99 en sus respectivas áreas.

Taller de Base de Datos

David Sol

En el contexto del ENC '99 se celebró el taller de Base de Datos, el cual comprendió la realización de tres actividades: la presentación de trabajos, la ponencia magistral del Dr. Ramez Elmasri y una discusión inicial para formar un grupo de interés en el área. Diez trabajos fueron presentados. En el grupo hubo propuestas de cuatro países: Chile, España, Francia y México. El taller se llevó a cabo con entusiasmo, y cada una de las presentaciones generó una gran participación del público; podemos estimar un promedio de 20 participantes por ponencia. Del total de participantes, alrededor una cuarta parte venía del sector empresarial y un 75% del sector académico, lo que permitió tener un grupo heterogéneo que aportó diferentes puntos de vista a los trabajos presentados.

Pudimos notar que el interés por las Bases de Datos como tema de investigación ha disminuido en el contexto nacional. Los participantes de España y Chile comentaron que en sus respectivos países se presenta un fenómeno similar: los grupos de investigadores en Bases de Datos son muy pequeños.

Los trabajos que se presentaron cubrieron cuatro áreas: Bases de Datos Distribuidas, Bases de Datos Geográficas, Modelado de Datos y Recuperación de Información. En la sesión de introducción hubo una presentación del responsable del área de Bases de Datos de GEDASNA, empresa de consultoría en software con unos 550 trabajadores, quien describió el estado actual de uso de las Bases de Datos en las empresas.

Podemos afirmar con agrado que gran parte de los trabajos presentados tienen una orientación para cubrir las necesidades actuales del sector empresarial. Asimismo, los comentarios de los investigadores nos permitieron también percatarnos de la calidad científica de las ponencias.

Hanna Oktaba

Taller de Ingeniería de software

El objetivo del taller de Ingeniería de Software (TIS) fue crear un foro para la presentación de la investigación en IS realizada en México, así como los avances de proyectos de tesis de doctorado, maestría y licenciatura, y la vinculación con la industria del software en México.

De los 24 trabajos, recibidos por nuestro taller y sometidos al proceso de evaluación realizado por 11 investigadores nacionales, se seleccionaron 19. De ellos la mayoría provenía del CIC-IPN (5), de la UNAM (4) y del CENIDET (2). El ITESM-Morelos, el CICESE, el IIE y el IT

Chihuahua II presentaron un trabajo cada uno, y por primera vez, contamos con trabajos de Cuba, España, Estados Unidos, Canadá y Brasil. Dos de los trabajos ofrecieron el avance de los proyectos de doctorado y varios más fueron resultados de tesis de maestría.

No se logró contar con la participación de la industria del desarrollo de software entre los ponentes aunque hubo una pequeña presencia entre los asistentes.

El tutorial y la conferencia magistral del TIS fueron ofrecidos por Kevin Dye, consultor experimentado, quien presentó una técnica para el trabajo colaborativo en el diseño de sistemas complejos, que tuvo mucho éxito, sobre todo entre los alumnos del posgrado de Ciencia e Ingeniería de la Computación de la UNAM, a tal grado que el expositor prolongó su estancia en México para probarla con un caso práctico de la definición de los requerimientos de los proyectos de software en dicho posgrado.

El tutorial se realizó con la asistencia de alrededor de 30 personas, mientras que las secciones del taller contaron en promedio con 20 asistentes.

Los comentarios críticos recibidos al final del taller se pueden resumir en los siguientes puntos:

1. Hubo poco tiempo para la presentación de trabajos (20 min.), situación que dificultó la discusión de los interesados en el tema.

2. Faltaron espacios sociales (coctel de bienvenida, cena del primer día) para poder relacionar de manera informal a los asistentes desde el inicio.

3. Faltó la información sobre el programa de cada taller en las puertas de los salones, así como la información oportuna sobre las actividades del evento.

4. El espacio físico y la realización de tantos talleres en paralelo no ayudó en el aprovechamiento del evento por parte de los asistentes.

5. Hubo demasiadas conferencias magistrales que quitaron tiempo para organizar mesas de discusión en los talleres.

En contrapartida, surgieron propuestas para mejorar los próximos eventos; a saber:

1. Agrupar los talleres en 2-3 grandes ramos, para que la gente no tenga que correr de uno al otro.

2. Realizar eventos sociales al inicio.

3. Incluir sesiones informales de café con mesas temáticas.

4. Invitar a gente de empresas.

5. Designar a un coordinador responsable del grupo de coordinadores de cada taller y no esperar hasta que alguien lo haga de manera espontánea.

6. Tener más atenciones para con los invitados por parte de los organizadores (reconocimientos, regalos, paseos, etc.)

Eventos de interés...

MICAI-2000

Conferencia Internacional Mexicana de Inteligencia Artificial.

Del 10 al 14 de abril del 2000.

Información:

<http://www-cia.mty.itesm.mx/micai2000>

PAAM 2000

The fifth International Conference and Exhibition on the practical application of Intelligent Agents and Multi-agent Technology.

Del 10 al 12 de abril del 2000.

Información:

<http://www.practical-applications.co.uk/PAAM2000/>

PA EXPO 2000

Información:

<http://www.practical-applications.co.uk/Expo2000/>

PACLIP 2000

Practical Application of Constraint Technology and Logic Programming,

Del 10 al 12 de abril del 2000.

Información:

<http://www.practical-applications.co.uk/>

PAKeM 2000

Practical Application of Knowledge Management, a celebrarse del 12 al 14 de abril del 2000.

<http://www.practical-applications.co.uk/>

PA Java 2000

Practical Application of Java, a celebrarse del 12 al 14 de abril del 2000.

Información:

<http://www.practical-applications.co.uk/>

Taller de Interacción Mediada por Computadora

Alfredo Sánchez

El Taller de Interacción Mediada por Computadora reunió a un importante número de participantes, observando un notable crecimiento en relación a su primera edición, realizada en 1997.

Se recibieron un total de 23 trabajos (casi un 300% más que en 1997) provenientes de 13 instituciones de educación superior o centros de investigación ubicados en ocho estados de la República (Baja California, Jalisco, Estado de México, Michoacán, Morelos, Puebla, Veracruz, Tamaulipas) y en el Distrito Federal. Se recibieron también trabajos de Argentina y España. El comité técnico fue integrado por investigadores de áreas como Interacción Humano-Computadora, Sistemas para Trabajo Cooperativo y Sistemas Multimediales, quienes se desempeñan en instituciones nacionales y extranjeras, incluyendo UNAM, UDLA-P, INAOE, CICESE, Carnegie Mellon University, Universidad Nacional de la Plata (Argentina), Universidad de Puerto Rico, Universidad Católica de Río de Janeiro y Universidad de Chile. En total, se aceptaron 12 trabajos.

El programa del Taller (<http://ict.udlap.mx/imc/enc99>) incluyó un curso tutorial sobre Usabilidad en Web, la presentación de trabajos arbitrados, y dos ponencias magistrales. Para la presentación de trabajos se organizaron cinco sesiones técnicas con los temas: Prototipos de CSCW, Interfaces Multimodales, Interfaces en 3D, Técnicas de Evaluación, y Agentes y Herramientas de Desarrollo. La participación del público fue excelente.

La primera ponencia magistral fue impartida por el Jaime Sánchez Ilabaca, profesor del departamento de Ciencias de la Computación de la Universidad de Chile, quien se enfocó a los aspectos principales del diseño de la interacción entre usuarios y ambientes computacionales. Enfatizó en principios probados de usabilidad y la necesidad de considerar a WWW como un medio distinto a otros ya existentes y, por tanto, con requerimientos particulares para el diseñador. La segunda ponencia magistral estuvo a cargo de Judy Brewer, directora de la Oficina Internacional de la Iniciativa de Accesibilidad al WWW (WAI, dentro de WWW Consortium). Ella presentó los esfuerzos realizados por WAI para desarrollar y fomentar el uso de guías que faciliten el acceso a la información a todo tipo de usuarios. Estas guías (<http://w3c.org/WAI>) se han convertido ya en una recomendación oficial de WWW Consortium, y aunque enfatizan el acceso a personas con discapacidades, se traducen en beneficios para todos los usuarios.

Se invitó a todos los asistentes a integrarse al Grupo de Interés en Interacción Mediada por Computadora de la Sociedad Mexicana de Ciencias de la Computación (<http://ict.udlap.mx/imc>). Se presentaron los objetivos del grupo, principalmente el de crear una comunidad mexicana en el área y coordinar los esfuerzos de investigación y desarrollo que se realizan en diferentes instituciones en áreas como Interacción Humano-Computadora, Sistemas para Trabajo Cooperativo (CSCW) y sistemas de multimedia e hipermedios.

Se presentó también el Capítulo CHI-México del Grupo de Interés en Interacción Humano Computadora de la ACM (<http://www.acm.org/chapters/chimex/>), el cual representa la faceta internacional de la comunidad mexicana en esta área. Cabe hacer notar que ACM CHI-México apoyó económicamente la realización del taller, por lo cual deseamos expresar nuestro agradecimiento.

Raúl Jacinto

Taller de Sistemas Distribuidos y Paralelos

Para la realización del Taller de Sistemas Distribuidos y Paralelos fueron presentados un total de 32 trabajos, de los cuales el comité técnico respectivo aceptó 22, aunque durante el evento únicamente se presentaron 16.

En este caso, la plática magistral —misma versó sobre Real Time Distributed Systems— estuvo a cargo de Ragunathan (Raj) Rajkumar.

Es importante señalar que, con la ayuda de profesores de la Universidad Tecnológica de la Mixteca, se levantó la minuta de la reunión y se cuenta con un medio electrónico de registro y de intercambio de información (<http://www.utm.mx/Z/SD>).

Concurso interno de Programación ACM

Ganadores del Concurso Interno de Programación de la ACM:

PRIMER LUGAR:

Nombre: *Waxtron*

Integrantes:

Juan Fco. Cardona Amozurrutia

Omar Ponce Carrillo

Ramón Argelles Garca

Suplente: _____

Coach: Dr. Mauricio Osorio Galindo

SEGUNDO LUGAR:

Nombre: *Novatos*

Integrantes:

Enrique Corona

Luis Ángel Montiel

Juan Antonio Navarro

Suplente: _____

Coach: Dr. Oleg Starostenko

TERCER LUGAR:

Nombre: *Los Intocables*

Integrantes:

Hugo Sánchez Ricardez

Reynolds Bonifaz Solrzano

Erick Lpez Guajardo

Suplente: Alejandro Dillanes Luna

Coach: Dr. Antonio Aguilera

ATTE.

Mesa Directiva 1999-2000
Java en Serio 5.0

Taller de Lógica y Computación

Raúl Monroy

Este taller incluyó 14 artículos, escritos por 21 investigadores afiliados, en total, a más de diez instituciones de educación superior y/o centros de investigación del país y otros. Cinco fueron los grandes temas estudiados: razonamiento automático, complejidad algorítmica, agentes y métodos de demostración y programación lógica, indicando que aún hay mucha dispersión en las líneas de investigación de la comunidad en el área. Comparado con el ENC'97, la autoría mostró un incremento en la participación femenina: casi el 30%. Las instituciones de afiliación con el índice más alto de contribución son, en este orden, la UNAM, el IPN, UDLA-P, y el campus Estado de México del ITESM.

En promedio, la audiencia por artículo fue de 25 personas, la mayoría investigadores, aunque hubo una presencia constante de estudiantes provenientes de la Universidad Autónoma de Hidalgo, la Universidad la Salle, etc. Entre la asistencia, destaca la participación de María del Pilar Pozos, quien recibió el financiamiento del comité organizador del ENC para cubrir los costos de transporte, inscripción, etc. María del Pilar es estudiante en el programa de doctorado del ONERA-Toulouse, Francia.

Asociados al taller de lógica y computación, se dictaron un tutorial y una conferencia magistral, ambas por el Dr. David Pearce, afiliado al DFKI, de las siglas en alemán de Centro Alemán de Investigación en Inteligencia Artificial. En estos eventos, el tema central lo constituyó la relación entre la lógica de equilibrio y el razonamiento no monótono.

Más de diez investigadores, algunos afiliados a instituciones extranjeras, realizaron la revisión técnica. Cada artículo fue revisado al menos por dos árbitros, rechazándose más de un 20% de los artículos recibidos. La mayoría de los contribuidores jóvenes tuvo la oportunidad de recibir calificación internacional, lo que seguramente repercutirá en la adecuación de estos jóvenes a foros de esa índole. El taller fue coordinado por Atocha Aliseda, del Instituto de Investigaciones Filosóficas de la UNAM, y Raúl Monroy del campus Estado de México del ITESM.

Isaac Rudomín

Taller de Graficación, Visión y Robótica

En este taller participaron 37 investigadores; se presentaron 19 artículos. Como referencia, la afiliación asociada a tal producción científica es principalmente universidades o centros de investigación, en el país y fuera de éste, incluyendo los campus Estado de México, Morelos y Monterrey del ITESM, la UNAM, el CIMAT, la UAM-Atzacapotzalco, el Instituto de Investigaciones Eléctricas, el CIC-IPN, el ITESO, el INAOE, la UDLA, el Imperial College (UK), University of Florida (USA), University of New Brunswick (Canadá).

El taller de Graficación Visión y Robótica fue coordinado por Isaac Rudomín del ITESM-CEM; Ana Luisa Solís, Víctor García Garduño y Boris Escalante Ramírez de la UNAM, y Carlos Avilés de la UAM-Atzacapotzalco. La revisión técnica quedó en manos de más de 15 investigadores. Cada artículo fue revisado al menos por dos árbitros, rechazándose más de un 40% de los trabajos recibidos, lográndose de esta manera un evento de alta calidad académica.

Taller de Aprendizaje, Búsqueda, Clasificación y Optimización con Enfoques Heurísticos

Eduardo Morales

Los responsables del taller fueron los doctores: Edgar Nelson Sánchez Camperos (CINVESTAV Guadalajara), Raúl R. Leal Ascencio (ITESO, Guadalajara), Luis Enrique Sucar Succar (ITESM Morelos) y Eduardo Morales Manzanares (ITESM Morelos).

Se recibieron 29 artículos, cada uno de los cuales fue evaluado en promedio por dos personas y sólo se aceptaron artículos cuyo dictamen fuera aceptado o marginalmente aceptado. El número final de artículos aceptados fue de 17, lo cual arroja un porcentaje de aceptación poco menor al 60%. La primera implicación de esto fue que la calidad de los artículos del taller fue superior con respecto a hace dos años; además, se tuvo mayor participación, en cuanto a artículos sometidos, y mayor diversidad de instituciones.

El domingo 12 de septiembre, el Dr. Ivan Bratko de la Universidad de Lubiana y del Josef Stefan Institute de Eslovenia presentó su tutorial titulado: "Inducing Intermediate Concepts in Machine Learning". También el Dr. Bratko impartió la primera conferencia magistral titulada: "Capturing Operator's Skill by Machine Learning". El lunes 13 se presentaron siete artículos relacionados con Aprendizaje por Refuerzo, Algoritmos Genéticos, Búsqueda Tabú y Recocido Simulado. El martes 14 se presentaron seis artículos sobre Redes Neuronales y Optimización. Finalmente, el 15 de septiembre se presentaron tres artículos sobre Búsqueda y Planificación, y se llevó a cabo un panel de discusión sobre técnicas de aprendizaje, desarrollo, comparación y perspectivas. El panel lo integraron Carlos Coello, L. Enrique Sucar, Raúl Rojas y Eduardo Morales.

XXXII Congreso Anual de la Sociedad Matemática Mexicana

En Guadalajara, Jal., del 10 al 15 de octubre pasado se celebró la más reciente edición del Congreso Anual de la Sociedad Matemática Mexicana, con más de 400 conferencias, expertos de distintas áreas de las matemáticas y temas relacionados. Uno de ellos, evidentemente, el de Ciencias de la Computación, donde hubieron un total de 20 conferencias de distintos tipos: divulgación, investigación, minicursos y reportes de tesis.

Los temas de computación involucrados fueron criptografía, semántica del lenguaje, algoritmos (cuánticos, paralelos, de reasignación de tareas), autómatas celulares, robótica, redes neuronales y redes funcionales.

La asistencia a las pláticas fue variable pero en promedio habían alrededor de 30 participantes por charla. Algunas pláticas tuvieron un gran éxito, como fueron por ejemplo, "Firmas digitales" de Julia Izta-cihuatl Muñoz Guerra y "Modelo matemático de un robot móvil" de Héctor Simón Vargas Martínez, como prueba de que ambos temas actualmente despiertan la curiosidad no sólo de los matemáticos sino del público en general (también pudo haber sido que los conferencistas eran muy populares).

Understanding Intelligence Rolf Pfeifer y Christian Scheier MIT Press



A mediados de los ochentas los investigadores en inteligencia artificial, ciencias de la computación, ciencias cognitivas y psicología se dieron cuenta de que la idea que se tenía de las computadoras como máquinas inteligentes era desafortunada o poco apropiada. El cerebro no "ejecuta programas"; hace cosas completamente diferentes. La teoría de la evolución dice que el cerebro humano ha evolucionado no para hacer demostraciones matemáticas sino para controlar nuestro comportamiento y para garantizar nuestra supervivencia. Muchos investigadores actualmente parecen estar de acuerdo en que la inteligencia se manifiesta como comportamiento, y por tanto es el comportamiento lo que debemos entender y estudiar. Así, se ha desarrollado toda una nueva área de investigación llamada inteligencia artificial, basada en el comportamiento.

Este libro proporciona una introducción sistemática a esta nueva forma de pensar. Después de presentar conceptos como las arquitecturas de subsunción, los vehículos de Braitenberg, la robótica evolutiva, la vida artificial, la auto-organización y el aprendizaje, los autores introducen un conjunto de principios y un marco teórico coherente para el estudio de sistemas inteligentes tanto naturales como artificiales y para el estudio de agentes autónomos. Este marco teórico está basado en una metodología sintética cuyo objetivo es el entendimiento mediante el diseño y la construcción.

Puedes obtener más información sobre esta obra o adquirirla en:
<http://www.ifi.unizh.ch/~pfeifer/undint.html>



José Cuenca, amigo de México

Hace unos meses y después de una larga afición, murió José Cuenca Bartolomé, quien fuera durante muchos años el presidente de la Asociación Española para la Inteligencia Artificial, y últimamente el director del departamento de Informática de la Universidad Politécnica de Madrid.

Formado como ingeniero civil (“camino canales y puertos”) había tenido una carrera distinguida como funcionario público y consultor independiente. A mediados de los ochenta se metió a académico transformándose en un maestro apreciado y prolífico.

Publicó dos libros que resultaron sumamente populares, uno sobre lógica y otro de sistemas expertos, e innumerables artículos. Fue un convencido promotor del uso de las técnicas de inteligencia artificial en conjunción con sistemas informáticos convencionales, siendo sus proyectos de control de tráfico y de control de inundaciones fluviales, trabajos de gran complejidad y notable éxito en la práctica. También fue un infatigable instigador del entendimiento internacional.

Fue Pepe un hombre bueno, generoso y lúcido. En México le debemos mucho. Vino, por primera vez, en 1985. Vino a dar un curso, pero se llevó un proyecto que habría de concretar a lo largo de una década: la colaboración en inteligencia artificial entre México y España, primero, y con Iberoamérica después. Hoy hay que recordar, entre otras cosas, que si no fuera por su fraternal y pertinaz insistencia, no habría una Sociedad Mexicana de Inteligencia Artificial, pues fue a través de su mirada generosa que esa incipiente comunidad tomó forma para poder dialogar con la contraparte española, en 1986. Para 1988 se convocó el primer Congreso Iberoamericano en Inteligencia Artificial (IBERAMIA) en el que ya participaba Portugal; mismo que cada dos años se ha venido realizando, con la participación de cada vez más países, hasta ahora bajo la prudente tutela de José Cuenca. En 1988, también, se firmó un convenio (CONACyT-CSIC) gracias al cual, año con año, ha habido un intercambio de investigadores entre México y España.

Es ley de vida morir, los hombres buenos, generosos y lúcidos como Pepe perviven, sin embargo, en sus obras. Sus discípulos y sus amigos, quienes lo conocimos, recordamos con gratitud y afecto el discreto y agudo actuar de este caballero español y amigo de nuestro país. Descanse en Paz.

Pablo Noriega B.V.

Convocatorias

CEC 2000

Congress on Evolutionary Computation

Son bienvenidas contribuciones que contengan avances en la teoría, práctica y aplicación de todas las técnicas evolutivas.

Información:

<http://pcgipseca.cee.hw.ac.uk/cec2000/>

Fecha límite para entrega de trabajos: 1º de febrero del 2000.

The 7th International Conference on Machine Learning

Del 29 de junio al 2 de julio del 2000.

Información:

<http://www-csli.stanford.edu/icml2k/>

Fecha límite para someter trabajos: 24 de enero del 2000.

Enseguida, el mensaje del Dr. Alfredo Sánchez, Presidente de la nueva mesa directiva de la Sociedad Mexicana de Ciencias de la Computación, 1999-2001, durante el acto de toma de protesta, celebrado el día 15 de septiembre pasado.

Convocatorias

CL 2000

First International Conference on
Computational Logic
Del 24 al 28 de julio del 2000

Información:

<http://www.doc.ic.ac.uk/c12000>

Fecha límite para entrega de trabajos:
1º de febrero del 2000.

COCOON 2000

6th Annual International
Computing and Combinatorics
Conference
Del 24 al 28 de julio del 2000.

Información:

cocoon2000@cse.unsw.edu.au

Fecha límite para entrega de trabajos:
1º de febrero del 2000.

KES 2000

4th International Conference on Knowledge-based Intelligent Engineering Systems and Allied
Techologies, a celebrarse del 30 de agosto a 1º de septiembre del 2000

Información: <http://www.eng.bton.ac.uk/eee/research/kes2000>

Fecha límite para entrega de trabajos 1º de marzo del 2000.

Hola a todos:

Saludo a Cristina Loyo y a todos los integrantes de la mesa directiva saliente:

Con estas breves palabras quiero expresar el reconocimiento y el agradecimiento de la comunidad de computación de México por su gran esfuerzo y grandes logros durante estos primeros años de la Sociedad Mexicana de Ciencias de la Computación.

Consolidar lo que ustedes nos han ayudado a lograr y contribuir a que esta comunidad siga desarrollándose constituyen retos enormes.

El Consejo Directivo entrante toma estos retos con entusiasmo. Contamos en esta empresa con la voluntad de participación de todos y cada uno de los miembros de la SMCC.

A los integrantes del Consejo Directivo saliente les decimos "hola", ni siquiera "hasta luego", ya que sabemos que contaremos permanentemente con su apoyo y experiencia.

A la SMCC en pleno le expresamos nuestro compromiso de hacer el mejor esfuerzo para contribuir a alcanzar los objetivos de esta sociedad.

¡Mil gracias!

Entrevista con Vladimir Estivill-Castro

Nombre: Vladimir Estivill-Castro

Fecha de nacimiento: 9/X/1963

Lugar de residencia: Newcastle, Australia

Puesto de trabajo: Senior Lecturer

Minibiografía: Investigador en Ciencias de la Computación y Consultor con amplia experiencia en Minería de Datos, Diseño y Análisis Orientado a Objetos, Aprendizaje Automático, Sistemas de Información Geográfica, Privacidad, Análisis Explorativo de Datos. Graduado de la Facultad de Ciencias de la UNAM en Matemáticas y con estudios doctorales de la Universidad de Waterloo en Canadá, ha participado en proyectos en Canadá, México, EU, Sudamérica y Australia. Actualmente encabeza un laboratorio de investigación en Minería de Datos y dirige varios trabajos de posgrado.



SMCC: Descríbenos tu trabajo cotidiano...

VE-C: Creo que me dedico a hacer ciencia y a la educación superior. Lo primero involucra mucha interacción vía e-mail con colaboradores. La experimentación con los algoritmos que diseño, incluye programar, pero también leer muchos artículos técnicos y científicos. En cuanto a la educación, doy clases en una universidad, superviso varias tesis y colaboro en las definiciones de los planes curriculares. También viajo frecuentemente a congresos como ponente y disfruto al analizar las presentaciones de otros colegas.

SMCC: ¿Cómo se relaciona tu trabajo con la vida práctica?

VE-C: Cada día es más claro que la economía es más compleja y que el desempleo es más marcado en sectores que no tienen acceso a la educación. Y por educación no me refiero a un entrenamiento técnico, sino a una formación crítica y profesional. También está relacionado con el avance tecnológico, que afecta la vida cotidiana. Además, frecuentemente hago consultorías a empresas en sus proyectos de Ingeniería de Software, en especial Orientación a Objetos y es mi satisfacción ver su impacto en la forma de construir sistemas.

SMCC: ¿Cuál es la situación de tu especialidad en el mundo?

VE-C: Como académico en el área de ciencias de la computación, creo que es muy buen momento. En Australia, los más de 30 Departamentos de Ciencias de la Computación tienen por lo menos una plaza disponible, y la situación es algo similar en muchas otras regiones. Sin embargo, a nivel mundial, el apoyo a la ciencia, a las universidades y a la educación está un poco frágil o no tan constante como antes de las recesiones de principios de los 90. En general, es triste que globalmente los gobiernos adopten una postura tendiente a considerar a la educación y a la investigación como un costo y no como una inversión.

SMCC: ¿Y en México?

VE-C: De México ahora sé muy poco. Creo que hay gran talento, muy buenos académicos y científicos, pero me resulta incomprensible la telaraña de intereses políticos, económicos, sociales que afectan a las instituciones públicas de educación y fomento a la ciencia y la cultura. México no tiene un sector privado extremadamente afluente que esté en condiciones de grandes inversiones en ciencia fundamental y educación. El país depende, en gran medida de un apoyo público para el desarrollo y fomento de la investigación, la cultura y la educación. Los sazones usan el término "montaña rusa", para algo que tiene grandes bajadas y grandes subidas. Como educador y científico, en México siempre me sentí en la montaña rusa, a veces con viento en popa y a toda vela, y a veces en la más oscura niebla.

SMCC: ¿Cómo ves el futuro?

VE-C: El de la investigación en informática, más diverso, amplio y creciente. La computadora es una revolución mucho más profunda que la imprenta. Cosas como el dinero, los periódicos, los centros comerciales, y muchas otras cosas más están directamente afectados por las opciones que se abren con la informática. La globalización hace que la competencia esté a un brinco de pantalla en el www. En estas condiciones, México puede estar en una posición ventajosa, ya que una gran parte de su población es muy joven y puede adoptar el cambio sin problemas.

SMCC: ¿Que recomiendas en el caso de México?

VE-C: Tan a la distancia es difícil. Pero ya que me preguntan, diría que al menos creo que se necesita un serio compromiso de definir objetivos claros y concretos en educación y fomento a la ciencia que vayan más allá de los ciclos sexenales, la heterogeneidad de las responsabilidades federales

Sociedad Mexicana de ciencias de la computación

SOLICITUD DE INSCRIPCIÓN

en educación y fomento a la investigación científica. Al mismo tiempo se necesita desterrar el que cada proyecto nacional es la oportunidad de unos para que les haga justicia la Revolución. Una revisión muy profunda de los parámetros con los cuales se evalúa el desempeño científico y eliminar una serie de barreras legislativas y burocráticas que hacen ciertas reglas más flexibles para unos y más rígidas para otros. Mi ejemplo clásico es el SNI y las definiciones de instituciones públicas y privadas, como si el investigador realizase menos ciencia en México por hacerlo en una institución privada que en una pública. No sé qué tanto ha cambiado la estructura del SNI; como lo conocí, su estructura, tal vez su propia existencia e instrumentación, es altamente cuestionable. Sus misteriosas excepciones generan una serie de distorsiones que contribuyen a un ambiente errático. Debo admitir que tal vez peor sería el panorama sin el SNI, pero hay ciertos desequilibrios que de ninguna manera han sido resueltos por el SNI.

SMCC: ¿Cómo te decidiste a estudiar computación?

VE-C: Siempre pensé que hacer ciencia es algo muy constructivo. Creo que el interés por la ciencia vino de mis padres, principalmente mi madre, y tomo el interés por las máquinas y la ingeniería de mi padre. Debo admitir que en diferentes círculos la gente me ve con diferentes perfiles, a veces soy Matemático, a veces hago "Computer Science"; otros me ven como Ingeniero de Software, produciendo soluciones y aplicaciones a problemas prácticos y concretos. Me gusta navegar estas áreas, interactuar con matemáticos puros o ingenieros y complementar sus habilidades.

SMCC: ¿Tienes héroes en la vida real?

VE-C: Donald Knuth, alguien que 1973 escribió tan brillantes libros (ofreciendo recompensas por errores tipográficos). Knuth ha hecho tan bellas contribuciones a las matemáticas discretas y desarrolló TeX y LaTeX con una preocupación por la automatización de la labor editorial y merece mi más profunda admiración.

SMCC: Virtudes y defectos

VE-C: Ser muy trabajador, es la virtud y defecto que se me ocurre. Es virtud pues me ha dado muchos resultados, estabilidad económica, la posibilidad de visitar todos los continentes, etc. Defecto, porque los días sólo tienen 24 horas y hay un montón de cosas que hacer en esta vida en compañía de familiares y amigos.

SMCC: ¿Quieres agregar algo más?

VE-C: Estoy muy contento de ver que existe un genuino interés en grupos de investigadores mexicanos que trasciende las barreras institucionales y que ha hecho de la SMCC algo vivo, dinámico, y aportador al medio científico en computación en México. Gracias.

Socio Titular Numerario (cuota anual: \$400.00)

Socio Estudiante (cuota anual: \$100.00)

Institución Educativa: _____

País: _____

Grado al que aspira: _____

Socio Institucional (cuota anual: \$2,000.00)

Nombre completo: _____

Institución e empresa: _____

Tipo de Institución Pública Privada

Puesto que ocupa: _____

Marque el grupo de interés al que desee pertenecer:

• Aprendizaje

– Redes Neuronales

– Computación Evolutiva

• Interfaces Humano-Computadora

• Lógica y Computación

• Visión, Robótica y Graficación

• Ingeniería de Software

• Bases de Datos

• Sistemas Distribuidos y Paralelos

Datos personales

Calle y número: _____

Colonia y ciudad: _____

C.P. y Estado: _____

País: _____

Teléfono y fax: _____

Dirección electrónica: _____

RFC: _____

Datos de la empresa o institución

Calle y número: _____

Colonia y ciudad: _____

C.P. y Estado: _____

País: _____

Teléfono y fax: _____

Dirección electrónica: _____

RFC: _____

Sociedad Mexicana de Ciencias de la Computación

Casa No. 24 UDLA-P. Cholula, Pue. C.P. 72820.

Tel.: (52) (2) 2-29-20-02.

URL: www.lania.mx/~smcc e-mail:

smcc@xalapa.lania.mx

Impacto Social de las Tecnologías de la Información Simposio

El 15 de octubre pasado concluyó el **Simposio Latinoamericano y del Caribe: Las Tecnologías de la Información en la sociedad. Uso e impacto presente y futuro**, organizado por el INEGI y el Programa intergubernamental de Informática de la UNESCO.

Las temáticas principales fueron: uso e impacto de las tecnologías de la información en la sociedad, y políticas de informática e información.

Se discutieron las nuevas formas de brindar servicios públicos como salud, turismo y seguridad; la importancia de los servicios públicos de información a los ciudadanos; el uso de tecnologías de información en el poder legislativo y la informatización de archivos. Asimismo, se abordaron temas sobre el comercio electrónico, el intercambio electrónico de datos, los medios de pago electrónico, trabajo a distancia y nuevos mercados laborales. Se trataron también tópicos sobre las redes de servicios comunitarios, la difusión de la informática entre los ciudadanos y la evolución de Internet en América Latina y el Caribe. Otros temas analizados en este escenario fueron: la protección de datos personales, delitos y seguridad en informática, soberanía e innovación en las redes globales, así como programas regionales e internacionales, financiamiento del sector público, la actividad del sector empresarial y comercial y lo referente a bancos regionales y nacionales.

¿Quién es quién... en la nueva mesa directiva de la SMCC

Presidente — Alfredo Sánchez



Alfredo Sánchez es director de la Biblioteca de la UDLA-P desde diciembre de 1998. Investigador Nacional nivel I del SNI, obtuvo los títulos de maestría (1993) y doctorado (1996) en computación, ambos de la Universidad de Texas A&M. Desde

1996 ha sido profesor en el Departamento de Ingeniería en Sistemas Computacionales de la UDLA-P, en donde funda y dirige desde enero de 1997 el Laboratorio de Tecnologías Interactivas y Cooperativas (ICT, <http://ict.udlap.mx>). Entre los proyectos de este laboratorio se encuentran la Biblioteca Digital Florística (<http://ict.udlap.mx/projects/fdl>), un proyecto colaborativo financiado por NSF, y los desarrollados en forma conjunta con la Biblioteca de la UDLA para dotar a sus usuarios de servicios personalizados. En otoño de 1997, ICT se convierte en uno de tres laboratorios que integran el Centro de Investigación en Tecnologías de Información y Automatización (CENTIA). Desde 1998 coordinada a un grupo de investigadores de instituciones mexicanas para definir un Programa Nacional de Bibliotecas Digitales. Con el apoyo de NSF y CONACyT, organizó un taller México-EU sobre Bibliotecas Digitales (<http://ict.udlap.mx/dl/workshop99>). Coordina el Grupo de Interés en Interacción Mediada por Computadora de la SMCC. Ha sido miembro de comités técnicos de congresos internacionales como los de Bibliotecas Digitales e Interacción Humano Computadora de ACM. Es miembro del Comité Académico de Computación e Informática del Ceneval y árbitro de proyectos para el Comité de Ingeniería Eléctrica, Ciencias de la Computación y Matemáticas del CONACyT. Es miembro de la Sociedad de Honor Phi Beta Delta por logros internacionales desde 1997, y de la Sociedad de Honor Phi Kappa Phi por logros académicos desde 1996.

Convocatorias

ASA/MA 2000

Second International Symposium on Agent Systems and Applications together with Fourth International Symposium on Mobile Agents
Del 13 al 15 de septiembre del 2000.

Información:

<http://www.inf.ethz.ch/ASA-MA/>

Fecha límite para entrega de trabajos: 3 marzo del 2000

ICEC 2000

International Conference on Electronic Commerce.
Del 21 al 24 de agosto del 2000

Información:

<http://icec.net>

¿Quién es quién... en la nueva mesa directiva de la SMCC

Vicepresidente

Carlos Coello



Carlos Artemio Coello Coello es originario de Tonalá, Chiapas. Cursó estudios de licenciatura en ingeniería civil en la Universidad Autónoma de Chiapas y posteriormente obtuvo una

beca de la SEP, para cursar estudios de maestría y doctorado en la Tulane University, en Nueva Orleans, Louisiana, de donde se graduó en 1996. Posteriormente, fue profesor visitante en DePauw University (en Greencastle, Indiana) y Senior Research Fellow en el Engineering Design Centre de la Universidad de Plymouth (en Inglaterra). Es investigador titular en el Laboratorio Nacional de Informática Avanzada. Sus áreas de interés son: computación evolutiva (algoritmos genéticos, estrategias evolutivas, programación evolutiva y programación genética), inteligencia artificial, meta-heurísticas (por ejemplo, búsqueda tabú) y optimización.

Secretario — Manuel Romero



Manuel Romero Salcedo recibió el grado de doctor en Informática del Instituto Nacional Politécnico de Grenoble, Francia, en 1998 y el grado de Maestro en Ciencias de la Computación de la UNAM, en 1993. Trabaja como Investigador de tiempo completo en el Departamento de Ciencias de la Computación en el IIMAS, UNAM. Su áreas de interés son: Trabajo Colaborativo Asistido por Computadora, Edición Colaborativa, Trabajo Colaborativo en el Web, Sistemas Distribuidos Orientados a Objetos, Consistencia de la Información Distribuida y Tolerancia a Fallas y Trabajo Móvil. El Dr. Manuel Romero es miembro del SNI Nivel I y también de la Red de Desarrollo e Investigación en Informática (REDII) del CONACyT en la UNAM. El Dr. Romero ha estado involucrado en varios proyectos de investigación nacionales e internacionales, algunos de ellos vinculados con la industria, lo cual le ha permitido validar y evaluar sus resultados de investigación.

Vocal — Amalia Duch



Nació en Mexico D.F. el 28 de febrero de 1969. Estudió Matemáticas aplicadas en la Universidad Autónoma de Aguascalientes (UAA). Empezó a trabajar en Ciencias de la Computación al realizar su tesis de licenciatura en algoritmos genéticos en LANIA, en donde fue asistente de investigación. Realizó estudios de Doctorado en Computación Teórica

en la Universidad Politécnica de Cataluña, España, especializándose en diseño y análisis de estructuras de datos y algoritmos. Ha publicado algunos artículos relacionados con diseño y análisis de estructuras de datos multidimensionales, tema en el que desarrolla su tesis doctoral, próxima a terminar. Actualmente es profesora de medio tiempo en la UAA, en la que imparte clases de matemáticas discretas, análisis numérico y álgebra.

Vocal — Antonio Martínez

Nació en México DF el 12 de agosto de 1962. Ingeniero agrónomo, especialista en industrias agrícolas (Universidad Autónoma Chapingo). Trabajó cinco años en la industria de alimentos balanceados para animales, en las áreas de producción y control de calidad; ahí vislumbró la necesidad de adquirir mayores conocimientos de estadística, por lo que realizó una maestría en el Colegio de Postgraduados. Entonces se le propuso formar parte de su plan de formación de profesores, lo cual consistía en hacer el doctorado para reincorporarse como profesor-investigador a su regreso. Así, realizó sus estudios doctorales en Ciencias de la Computación en la Universidad de Tulane en Nueva Orleans (1992). Se reincorporó al CP en 1998, como profesor-investigador adjunto de la especialidad de cómputo aplicado. Área de interés: intersección de las herramientas estadísticas aplicadas a la minería de bases de datos agrícolas.

Vocal — Eduardo Morales

El Dr. Eduardo Morales Manzanares nació el 4 de julio de 1961 en México, D.F. Doctor en Computación por el Turing Institute de la Universidad de Strathclyde (Glasgow, Reino Unido). Realizó una Maestría en Inteligencia Artificial en la Universidad de Edimburgo, también en el Reino Unido, y la Licenciatura en Ingeniería Física en la UAM. Ha trabajado como investigador visitante en el Electric Power Research Institute (EPRI), en Palo Alto, California. Trabajó como Investigador del Departamento de Simulación, del Instituto de Investigaciones Eléctricas y fue Consultor del proyecto ESPRIT "The Machine Learning Toolkit" de la Comunidad Económica Europea (CEE). Fue Investigador del Departamento de Sistemas de Información, también del Instituto de Investigaciones Eléctricas. Es profesor-investigador de la División de Ingeniería y Ciencias del ITESM - Campus Morelos. Ha impartido clases tanto a nivel maestría, como a nivel doctorado, en el área de Inteligencia Artificial. Ha participado en proyectos de investigación y desarrollo financiados por el Instituto de Investigaciones Eléctricas, la CFE, la CE, el EPRI, la Fundación Rockefeller y el CONACyT. Actualmente es responsable de un proyecto en Descubrimiento de Conocimiento en Bases de Datos. Cuenta con alrededor de 50 publicaciones en revistas, libros y congresos, tanto nacionales como internacionales, y es Investigador Nacional Nivel I del SNI.

Vocal — Carlos Avilés

Nació el 13 de marzo de 1966 en la Ciudad de México. Ingeniero en Electrónica con especialidad en Sistemas Digitales por la UAM. Realizó una Maestría en Procesamiento Digital de Señales, Imágenes y Voz en el Instituto Politécnico Nacional de Grenoble. Doctor por el Instituto Politécnico Nacional de Grenoble. En la UAM Azcapotzalco es Profesor-investigador Titular "A". El Dr. Avilés es autor y coautor de más de 15 artículos en revistas y congresos nacionales e internacionales, y es coautor de un libro. Áreas de interés: visión por computadora, procesamiento digital de imágenes y señales y estadística de orden superior. Candidato a investigador del SNI.

Vocal — Marcelo Mejía

Marcelo Mejía Olvera es Ingeniero en Biomédica por la UAM. Realizó una Maestría en Ciencias de la Computación en la UAM, en Ingeniería Eléctrica en la UNAM, y en Redes Informáticas en la Escuela Superior de Electricidad en Francia. En 1989 obtuvo el grado de Doctor en Informática de la Universidad de Rennes I, en Francia. Ha impartido cursos de licenciatura y posgrado en la UAM, el ITESM y el ITAM. Es Jefe del Departamento Académico de Computación del ITAM y Editor Técnico de la revista *Soluciones Avanzadas*.

Nueva Mesa de LyC

Estimado miembro del grupo de lógica y computación:

Te informamos que se registró una sola planilla para la nueva mesa del grupo de interés en lógica y computación de la SMCC. Por lo tanto, no fue necesario hacer elecciones y quedamos los propuestos inicialmente:

Atocha Aliseda	(atocha@filosoficas.unam.mx , http://www.filosoficas.unam.mx/~atocha)
Raul Monroy	(raulm@campus.cem.itesm.mx , http://research.cem.itesm.mx/raulm)
Alberto Oliart	(aoliart@xalapa.lania.mx , http://www.lania.mx/~aoliart)
Mauricio Osorio	(josorio@mail.udlap.mx , http://gente.pue.udlap.mx/~josorio)

Esta mesa estará en funciones hasta el ENC 2001. Esperamos contar con tu participación activa en este grupo, en dirección a consolidar nuestro grupo de lógica y computación en México.

Aprovechamos para darte nuestros datos y pedirte que actualices los tuyos. Si tienes otro correo o has cambiado de institución, por favor avísanos. Si conoces a otros interesados, por favor pídeles que se registren en el grupo. Esto puede hacerse a través de la página hogar del grupo. Nuestra misión inicial será la de depurar la lista de miembros y las páginas electrónicas con las que contamos. Te invitamos a visitarlas y a enviarnos tus sugerencias, quejas e ideas nuevas. En breve, pondremos los materiales del tutorial que el Dr. David Pearce impartió en el taller de lógica y computación celebrado en Pachuca en el marco del ENC '99.

Página del grupo: <http://research.cem.itesm.mx/~lyc>
Página del taller: <http://www.filosoficas.unam.mx/~lyc>
Lista grupo: lyc@campus.cem.itesm.mx Clave correo mesa: lyc@filosoficas.unam.mx

Atentamente

Atocha Aliseda, Coordinadora del Grupo de Interés en Lógica y Computación, SCC

Hallando los temas principales en un artículo en español

Adolfo Guzmán Arenas
Centro de Investigación en
Computación del I. P. N.

RESUMEN

Muchas operaciones sobre colecciones sistemáticas de datos pueden realizarse fácilmente por la computadora cuando los datos son numéricos. En cambio, cuando los datos están en documentos en español, organizados en secciones, párrafos y oraciones, no es posible que la computadora realice las operaciones anteriores. Puesto que gran parte del conocimiento humano está en textos en lenguaje natural, conviene descubrir métodos para realizar tales operaciones. Para ello, la máquina debe entender el texto, comprenderlo. Este pequeño artículo es un resumen de [2], y expone cómo analizar un documento escrito en lenguaje natural, y reconocer los temas principales de los que habla.

1. INTRODUCCIÓN Y OBJETIVOS

En el Centro de Investigación en Computación del Instituto Politécnico Nacional, el Laboratorio de Lenguaje Natural y Procesamiento de Texto está investigando el procesamiento inteligente de texto, necesario para llevar a cabo las labores ejemplificadas en el resumen. Una de tales labores, el hallazgo de los temas principales que un artículo en español aborda, ha sido resuelta y es el tema de esta ponencia.

1.1 La importancia de analizar español escrito

El análisis inteligente de texto permitirá que la computadora entienda los documentos escritos en lenguaje natural, por ejemplo, que haga resúmenes, encuentre tendencias, compare documentos (con respecto a un tema dado) y conteste preguntas no triviales.

Otra ventaja: constituiría para la máquina un nuevo método para aprender: sólo tendría que leer libros, los mismos libros que las personas utilizan. Para enseñarle Física, ya no sería necesario fabricar procedimientos y estructuras de datos que guardaran conceptos y fórmulas de esa ciencia. Nosotros avanzamos una parte pequeña en esta dirección, al poder saber qué temas o conceptos principales cubre un artículo.

1.2 Diferencia entre palabra y concepto

Una palabra o vocablo puede representar varios conceptos o acepciones. Los conceptos, en cambio (por definición) no son ambiguos:

ESTRELLA (palabra)

- estrella-astronómica (concepto) cuerpo sideral
- estrella-artista (concepto) persona famosa o destacada
- estrella-insignia (concepto) adorno

Para convertir las palabras en el concepto "correcto" hay que desambiguar: escoger para cada palabra su acepción (concepto) correcta, de varias posibles. Para desambiguar hay que utilizar información del contexto (del resto del artículo, por ejemplo, o de los temas de los que el artículo o la oración habla). Esto se denomina análisis semántico, ya que atañe al significado de cada palabra.

1.3 Los conceptos forman árboles

Es interesante notar que los conceptos están organizados en familias o jerarquías, como puede constatarse al hojear un Thesaurus. La relación más usada es la de "subconjunto", y los conceptos "hijos" denotan casos particulares o más específicos que sus padres. Tomemos un concepto que denota un objeto (sustantivo, sujeto) como ejemplo:

PRENDA DE VESTIR (concepto)

ZAPATO

- huarache
- sandalia
- mocasín
- bota
- alpargata
- pantufla

CAMISA

- camisa de manga corta
- camisa de manga larga

PANTALÓN

En este ejemplo, el concepto mocasín es un subconjunto del concepto zapato, lo que significa: todos los mocasines son zapatos, pero no todos los zapatos son mocasines. Hemos mostrado una pequeña parte del árbol de conceptos.

También los verbos (acciones, procesos) forman árbol, bajo la relación "subconjunto". Existen muchas otras relaciones ("parte de", "nacido en", "hijo de") que enriquecen el árbol del sentido común [Refs. 1 y 2].

En particular, el número de conceptos del conocimiento común es finito. Resulta, pues, posible formar los árboles de conceptos, o de una parte de ellos, y en esto se basa el algoritmo y programa, llamado Clasitex, que halla los temas en un artículo.

2. HALLANDO LOS TEMAS DE UN ARTÍCULO, INTENTO 1.

Como primer intento, podemos contar las palabras en el artículo. Aquéllas que más aparecen sin duda constituyen o denotan temas importantes. Por ejemplo, si aparece 17 veces la palabra "revolución", el artículo habla sobre revolución.

2.1 Problema: muchas palabras se refieren al mismo tema

Por ejemplo, un artículo sólo menciona “revolución” dos veces, pero menciona a Zapata, a Francisco Villa, a Flores Magón, ...

Por ejemplo, portero, defensa izquierdo, Atlante, Necaxa, Luis García, todos hablan (sugieren el tema) fútbol soccer.

Esto sugiere que fabriquemos el árbol de conceptos, y que reemplacemos las palabras en el artículo por los conceptos respectivos. Hecho esto, hallemos los conceptos más populares: ellos son los temas principales buscados. Al contar conceptos, debemos contar no sólo los conceptos “inmediatos” que una palabra denota (por ejemplo, la palabra “huarache” denota el concepto “huarache”), sino los conceptos hacia la raíz del árbol, que en el ejemplo serían zapato, prenda de vestir, etc. El método busca marcar o contar sobre el árbol de conceptos. Al finalizar el análisis del artículo, ciertas partes del árbol quedarán más marcadas (habrá contadores con mayores cantidades), y esas partes del árbol son los temas del artículo.

Problemas en el enfoque propuesto:

- Hay que hacer el árbol, manualmente. (Esto lo resolvemos haciendo en efecto una introducción manual).
- Mapear una palabra al concepto que denota no es fácil, pues es necesario desambiguar (Cf. §1.2). Si no lo hiciéramos, una palabra incrementaría el contador de cada uno de los conceptos que denota. (Esto se resuelve en Clasitex omitiendo la desambiguación, cosa que no introduce un incremento apreciable en la inexactitud, como veremos abajo).

2.2 Resumen del procedimiento propuesto

Halle los conceptos que más aparecen —esos son los temas. Para esto, observemos lo que sigue:

1. Contemos o miremos los conceptos, no las palabras.

Refinamiento: Cuente los conceptos sobre el árbol de conceptos.

Cuente los conceptos y sus generalizaciones (conceptos inmediatamente más generales)

- Cuando halle “huarache”, cuente también zapato, prenda de vestir, ...
En la versión actual de Clasitex, no se cuentan las generalizaciones, y no obstante los resultados obtenidos son aceptables. ♣
- Algunas partes del árbol se ennegrecerán más que otras. Las regiones negras del árbol señalan los temas del artículo.

Mientras más extenso es el artículo, mejor resulta este método.

2. Cada palabra “vota” por (aumenta el contador de) los conceptos que denota.

cuando hallamos canal aumentamos 1 al concepto canal-de-agua, y al concepto canal-de-comunicaciones.

3. Habrá votos equivocados. Empero, a la larga, los votos correctos prevalecen.

A) También es posible desambiguar en un segundo paso: las palabras que votaron más de una vez retiran todos excepto uno de sus votos (conservan el voto “correcto”), una vez encontrados los temas principales al final del paso 1. ♣ Estos votos que se retiran no deberían (salvo en contadas ocasiones) cambiar los temas principales encontrados.

B) En la práctica, hemos visto que este segundo paso no es necesario: la sola acumulación de los temas, tanto correctos como equivocados, hace que los correctos destaquen. Es como si, en vez de eliminar el ruido de una señal, acumulamos muchas de ellas, de modo que el ruido resulta reducido en raíz cuadrada de n , el número de señales agrupadas.

2.3 Hay conceptos denotados por una secuencia de más de una palabra en español

Por ejemplo,

- COCINA (una palabra) – denota el concepto “cocina”, habitación donde se confeccionan alimentos.
- NUEVO MÉXICO (dos palabras). Denota el concepto “Nuevo-México”, un estado de E.U.
- Benito Juárez (dos palabras). Denota al Benemérito de las Américas.
- Los Tres Mosqueteros (tres palabras). Denota un libro de Alejandro Dumas.
- Instituto Politécnico Nacional (tres palabras). Denota un famoso Instituto Tecnológico.

Lo que esto implica es que, al barrer el texto, debemos buscar no sólo palabras, sino pares, tríos, cuartetos, ... de palabras, en nuestro recorrido del texto, averiguando para cada una de estas palabras, pares, tríos, etc., si denotan algún concepto o no.

2.4 El árbol usado en Clasitex

Este último ejemplo nos permite emplear en Clasitex un árbol que usa, en vez de la relación “subconjunto” entre sus nodos (ver §1.2), la relación “sugiere el tema de”, o “vota por el tema ...”. Por ejemplo, “Benito Juárez” sugiere el tema “México”, el tema “Oaxaca”, y el tema “Historia de México”, aunque Benito Juárez es una persona, México es un país, etc.

2.4.1 Palabras sin sentido

Esta simplificación tiene otra consecuencia. Una

palabra como “tres” (que representa al concepto “3”, un número impar) o como “entonces” no sugiere ningún tema, no vota por ningún tema. A estas palabras las hemos llamado, un tanto inapropiadamente, de “sentido débil” o “sin sentido”, dando a entender que no contribuyen a dilucidar el tema del artículo.

3. PROCEDIMIENTO USADO EN CLASITEX

Describimos el algoritmo utilizado por Clasitex.

Recorremos con un apuntador el artículo en español (que está en un archivo), de izquierda a derecha.

1. (COMIENZO) Observa la secuencia o cadena de cuatro palabras señaladas por el apuntador. ¿Denotan algún o algunos conceptos?
Sí: aumenta en 1 el contador de cada uno de esos conceptos. Ve al paso (5).
No: Ve al paso siguiente (2).
2. Observa la secuencia o cadena de tres palabras señaladas por el apuntador. ¿Denotan algún o algunos conceptos?
Sí: aumenta en 1 el contador de cada uno de esos conceptos. Ve al paso (5).
No: Ve al paso siguiente (3).
3. *Ídem* la secuencia de dos palabras.
4. Observa la palabra señalada por el apuntador. ¿Denota algún o algunos conceptos?
Sí: aumenta en 1 el contador de cada uno de esos conceptos. Ve al paso (5).
No: ¿Es una palabra que no denota concepto alguno? Es decir, que sea palabra “sin sentido” según el §3.4.1.
Sí: ignórala. Ve al paso (5) siguiente.
No: Imprímela como “no sé qué significa esta palabra, o si no significa algo” (Hay que agregar estas palabras en alguna ocasión posterior al árbol de Clasitex, o al conjunto de palabras “sin sentido”). Ve al paso (5) siguiente.
5. Mueve el apuntador a la derecha de la(s) palabra(s) ya analizada(s). Y repite la operación o iteración: ve al paso 1 (COMIENZO). Cuando ya se haya analizado todo el texto (ya no podrá moverse el apuntador a la derecha), cuéntense los conceptos y repórtense los más numerosos como los temas del artículo.

En la práctica, se ha visto que el histograma de conceptos contiene una separación bien diferenciada entre los temas numerosos y los temas poco numerosos. Abajo veremos ejemplos.

3.1 Conclusiones y comentarios

Hemos restringido nuestro análisis de textos en español a artículos que aparecen en revistas y periódicos de circulación general, no especializada. Esto con el fin de limitar el número de conceptos especializados. El enfoque sirve para cualquier otro lenguaje natural.

En artículos cortos (México en el Mundial), Clasitex ofrece menos seguridad en sus temas: los histogramas contienen menos conceptos, y la diferencia entre los temas principales y los temas secundarios (y el ruido) es menos marcada.

El enfoque propuesto parece prometedor, pero hay que hacer muchos más experimentos para llegar a alguna conclusión definitiva. En particular, este enfoque requiere construir manualmente el árbol de conceptos.

La implementación actual está hecha usando el lenguaje de comandos (Shell) de Unix, y las utilerías awk, sed, etc. Esto permite modificaciones fáciles al programa, pero da una ejecución lenta (aproximadamente a la velocidad de lectura de una persona). Cuando se tenga una versión estable, reescribiremos los programas en C, guardando el árbol en estructuras en memoria, accedidas con índices. Emplearemos dispersión al azar (hashing) si se usare memoria secundaria.

Algunas mejoras a Clasitex aparecen marcadas con ♣ en el texto. Lo principal es aumentar considerablemente el árbol de conceptos.

3.2 Bibliografía

1. Douglas B. Lenat and R. V. Guha. *Building large knowledge-based systems*. Addison Wesley. 1990.
2. A. Guzmán. *Finding the main themes in a Spanish document*. To appear in *Expert Systems with Applications*, Elsevier, Jan. 1998.
3. A. Guzmán. “Hallando los temas principales en un artículo en español”. *Soluciones Avanzadas* 5, 47 (15 julio 97), págs. 58-63, y 5, 49 (15 sept. 97), págs. 66-71.

Suscríbete al

boletín
de la
Sociedad
mexicana
de ciencias
de la
computación

Suscripción
Anual:
\$100.00

Tel.: (52) (2) 2-29-20-02
URL: www.lania.mx/~smcc
e-mail: smcc@xalapa.lania.mx

¿Publicidad?



Anúnciate

en el boletín de la SMCC.

Tel.: (52) (2) 2-29-20-02
URL: www.lania.mx/~smcc
e-mail: smcc@xalapa.lania.mx