

REVISTA

BITS

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

de Ciencia
UNIVERSIDAD DE CHILE

Nº 5 / PRIMER SEMESTRE 2011

Gráfica / Computación Científica
Inteligentes / Computación
Colaborativos / Sistemas
Humano-Computador / Sistemas
Información / Interacción
de Datos / Sistemas de
Criptografía / Estructura
/ Lenguajes / Algoritmos /
Hardware / Redes / Software

Panorama de la Ciencia de la Computación en Chile

- **JOSÉ A. PINO:**
EL DCC ENTRE 1978-1983:
AUDENTES FORTUNA IUVAT
- **LEOPOLDO BERTOSSI:**
EN 20 AÑOS DE COMPUTACIÓN: UNA
VISIÓN MUY PERSONAL
- **PROGRAMAS DE POSGRADO CHILENOS
EN CIENCIAS DE LA COMPUTACIÓN E
INFORMÁTICA: UN DIAGNÓSTICO INICIAL**

Comité Editorial

Nelson Baloian, profesor.
Claudio Gutiérrez, profesor.
Alejandro Hevia, profesor.
Gonzalo Navarro, profesor.
Sergio Ochoa, profesor.

Editor General

Pablo Barceló

Editora Periodística

Ana G. Martínez

Periodista

Karin Riquelme

Diseño y Diagramación

Sociedad Publisiga Ltda.

Fotografías, gentileza de:

Gastón Carreño
Comunicaciones FCFM

Dirección

Departamento de Ciencias
de la Computación
Avda. Blanco Encalada 2120, 3º piso
Santiago, Chile.
837-0459 Santiago
www.dcc.uchile.cl
Teléfono: 56-2-9780652
Fax: 56-2-6895531
revista@dcc.uchile.cl

Revista BITS DE CIENCIA es una publicación del Departamento de Ciencias de la Computación de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile. La reproducción total o parcial de sus contenidos debe citar el nombre de la Revista y su Institución.

Revista Bits de Ciencia N° 5
ISSN 0718-8005 (versión impresa)

www.dcc.uchile.cl/revista
ISSN 0718-8013 (versión en línea)

CONTENIDOS

INVESTIGACIÓN DESTACADA

02 Minería de Repositorios de Software para ayudar a los desarrolladores

Romain Robbes

COMPUTACIÓN Y SOCIEDAD

08 El DCC entre 1978-1983: Audentes fortuna iuvat

José A. Pino

14 En 20 años de computación: una visión muy personal

Leopoldo Bertossi

PANORAMA DE LA CIENCIA DE LA COMPUTACIÓN EN CHILE

24 Programas de posgrado chilenos en Ciencia de la Computación e Informática: un diagnóstico inicial

Sergio F. Ochoa, Marcello Visconti

29 Panorama de la Ciencia de la Computación en Chile

Pablo Barceló

31 Hardware + Redes

Reinaldo Vallejos, NIC Chile Research Labs

33 Software + Lenguajes

Alex Bergel, Cecilia Bastarrica, Eric Tanter, Johan Fabry, Romain Robbes, SIGSE

40 Algoritmos + Criptografía + Estructura de Datos

CLCERT, Gonzalo Navarro, Jérémy Barbay, José Rafael Correa

44 Sistemas de Información

Ma. Andrea Rodríguez, Benjamin Bustos, Claudio Gutiérrez, Loreto Bravo, Marcelo Arenas, Mónica Caniupán, Mauricio Marín, Pablo Barceló

52 Sistemas Colaborativos + Interacción Humano Computador

CARL, Centro de Computación y Comunicación para la Construcción del Conocimiento

53 Computación Gráfica + Computación Científica

Ma. Cecilia Rivara, Fernando Rannou, Luis Salinas, Nancy Hitschfeld

58 Sistemas Inteligentes

Carlos Castro, Carlos Hernández, INCA, John Atkinson, NLCM, COMET, Mauricio Solar, GRIMA

SURVEYS

68 Contornos activos para segmentación en imágenes digitales

Jorge Jara

CONVERSACIONES

74 Entrevista a Andrea Rodríguez

Claudio Gutiérrez

78 Entrevista a Gonzalo Navarro

Claudio Gutiérrez

GRUPOS DE INVESTIGACIÓN

82 Gaming for Learning: interfaces interactivas para asistir el aprendizaje y la cognición

Jaime Sánchez

CONFERENCIAS

87 Alberto Mendelzon Workshop en Fundamentos de Bases de Datos (AMW 2011)

EDITORIAL



Los tres números anteriores de la Revista Bits de Ciencia fueron dedicados, consecutivamente, a los temas de Ciencia de la Web, Educación en Informática y TICs, y para cada uno de ellos la Revista analizaba su estado en Chile hoy. Todos estos temas son de alta importancia para el desarrollo del país, y sin duda también son de interés transversal para los distintos grupos cercanos a la Computación e Informática en Chile, que son los destinatarios naturales de nuestra Revista. Me refiero aquí a académicos, empresarios, desarrolladores, personeros de gobierno interesados en temas de tecnología de información, y lectores de otras áreas que tienen interés por nuestra disciplina.

Esta vez, sin embargo, hemos querido hacer un número un poco más académico, en un tema que de alguna forma es también un poco más cercano a nuestro corazón: el estado de la Ciencia de la Computación en nuestro país. Esto porque, a pesar de que nosotros los académicos de la especialidad vivimos inmersos en este ambiente -desarrollando investigación, generando proyectos, graduando alumnos de posgrado- muy pocas veces nos detenemos a observar cómo hacemos las cosas, quién conforma nuestra comunidad, cuántos alumnos estamos graduando etc. Y, sobre todo, muy pocas veces nos damos el tiempo de ver qué están haciendo nuestros pares en el país.

El presente número de la Revista trata de subsanar esto de alguna forma. Por supuesto, realizar un análisis exhaustivo del estado de la Ciencia de la Computación es una tarea titánica que escapa a nuestras competencias. Por tanto, no se puede pensar en ningún caso el presente número de la Bits como un trabajo monolítico que represente de una vez y para siempre cómo se hace Computación en Chile hoy.

Esta es la razón por la que hemos llamado a nuestro trabajo "Panorama de la Ciencia de la Computación en Chile". Lo que hemos pretendido es entregar ciertos esbozos y directrices de dónde estamos y qué es lo que hacemos.

Para ello invitamos a los integrantes de la comunidad de Computación del país a contarnos qué están haciendo, con quién trabajan, cómo se relacionan con sus pares chilenos, y dónde publican regularmente. Tal invitación ha tratado de ser lo más inclusiva posible, intentando no dejar a ningún investigador de experiencia fuera de ella. Es evidente que más de algún error hemos cometido, y como editor de la Revista me hago absolutamente responsable de ellos.

Como es usual, el tema central de la Revista es acompañado por varios artículos. En particular, incluimos la segunda parte de la historia del DCC, escrita por José A. Pino, una visión muy particular del desarrollo y estado de la Ciencia de la Computación en Chile, escrita por Leo Bertossi, y un análisis en profundidad de los programas de estudios de posgrado en Computación impartidos en Chile, escrito por Sergio Ochoa y Marcello Visconti. Además, incluimos nuestras secciones usuales de Investigación Destacada, Entrevistas y Grupos de Investigación. No está de más recordar que las opiniones y comentarios que se expresan en estos artículos, se emiten bajo exclusiva responsabilidad de sus autores.

Espero sinceramente que el presente número sea de interés no sólo de la comunidad académica, sino también de todo el resto de las personas interesadas en Computación e Informática en Chile.

Profesor Pablo Barceló
Editor Revista Bits de Ciencia

Minería de Repositorios de Software para ayudar a los desarrolladores

Uno de los primeros estudios empíricos de desarrollo de software fue el estudio de Lehman y Belady sobre el OS/360 de IBM en los años '70: Ese trabajo fue el origen de las leyes de Lehman sobre la evolución del software [Lehman1985] en las que se establece lo siguiente:

- “Los sistemas de software deben ser continuamente adaptados o se vuelven cada vez menos satisfactorios” (Ley 1).
- “Cuando un sistema de software está cambiando, su complejidad aumenta a menos que algún trabajo se realice para mantener o reducir su complejidad” (Ley 2).
- “La calidad de los sistemas de software parece estar disminuyendo a menos que sean rigurosamente mantenidos y adaptados a los cambios de entorno operativo” (Ley 7).

Desde entonces, las leyes de la evolución del software han sido verificadas una y otra vez.

El mantenimiento de software, es decir, la fase donde los desarrolladores reaccionan a los cambios de requisitos, es la fase más costosa del desarrollo, porque cambiar un sistema grande es extremadamente difícil. ¿Cómo se puede juzgar entonces el impacto de un cambio en un proyecto que ha vivido varios años, ha sido escrito por un equipo de desarrollo de gran tamaño (potencialmente distribuido), y tiene miles -si no millones- de líneas de código?

En este contexto, cada información adicional sobre el sistema que se mantiene es útil.

Para hacer frente a estos desafíos, los investigadores en el área de investigación llamada Mining Software Repositories (MSR) exploran los repositorios de información más precisos que tenemos sobre el



Romain Robbes

Profesor Asistente, DCC, Universidad de Chile. Doctor en Ciencias de la Computación, Universidad de Lugano, Suiza (2008); Diplome d'etudes approfondies, University of Caen, Francia (2003); Maitrise d'informatique, University of Caen, Francia (2002); DEUG and Licence, informatique (Bachelor), University of Caen, Francia (2001).
rrobbes@dcc.uchile.cl

desarrollo de software: los desarrolladores almacenarán discusiones de diseño, informes de problemas, y los cambios que realizan en herramientas especializadas, llamadas repositorios de software. Una vez recuperada – en un proceso que puede ser difícil - la información tiene un enorme valor para los desarrolladores, testadores, mantenedores, arquitectos y administradores.

En este artículo, en primer lugar, presentaré el tipo de información que se encuentra en dos tipos de repositorios de software: el sistema de control de versiones y el sistema de seguimiento de errores. Luego presentaré cinco enfoques que podrían aprovechar esta información para ayudar al proceso de desarrollo de software de varias maneras distintas.

Por último, voy a presentar algunos de los trabajos recientes en que he estado involucrado, que comparten un objetivo común: ¿Cómo podemos hacer más? La respuesta está en el uso de diferentes repositorios de software, es decir, repositorios de correo electrónico, super-repositorios y repositorios que contienen muchos más detalles sobre las actividades de los desarrolladores.

LOS REPOSITORIOS DE SOFTWARE

El sistema de control de versiones

El Software Configuration Management nace de la necesidad de coordinar el cambio, cuando varios desarrolladores están trabajando en el mismo sistema de software. ¿Cómo podemos asegurarnos de que los cambios de uno no se oponen a los de otra persona? Sin la disciplina adecuada, romper el código de otro desarrollador es extremadamente fácil. Del mismo modo, hay que encontrar una manera eficaz para revisar e integrar los cambios realizados por otra persona en la base de código.

El sistema de control de versiones se encarga de estas cuestiones.

Limitamos la discusión a los sistemas de control de versiones centralizados, como CVS y Subversion. Los sistemas de control de versiones distribuidos como Git o Mercurial funcionan de forma ligeramente diferente y plantean nuevos desafíos para los investigadores de MSR [Bird2009].

Un escenario típico es el siguiente: un programador se conecta al servidor de control de versiones, y recupera la versión más reciente del sistema (check out). Luego trabaja en sus tareas asignadas. En cualquier momento, es libre de hacer un commit, (check in), con sus cambios pendientes. A continuación, se reflejarán los cambios en el servidor; otros desarrolladores pueden integrar estos cambios la próxima vez que hacen un “check out” del sistema. Esta serie de cambios se llama una transacción o un commit. La Figura 1 muestra un “change log” que es una lista de commits. La información típica que se encuentra en un commit - que puede ser explotada posteriormente por enfoques MSR-, es la siguiente:

- El número de la transacción, que da una identidad única para referencia futura.
- El autor del commit (¿quién?).
- La fecha del commit (¿cuándo?).
- Un comentario sobre el commit: texto libre, idealmente sería una razón y una descripción de los cambios (¿por qué?).
- La lista de ficheros cambiados en el commit (¿cuáles?).

- Las líneas añadidas y eliminadas, por cada uno de los ficheros modificados (¿cómo?).

Como veremos a continuación, incluso con estos datos, aunque limitados de información, ya podemos ayudar eficazmente a un profesional.

El sistema de seguimiento de problemas

Más allá de tener varios desarrolladores, los sistemas grandes cuentan con muchas tareas a realizar, y muchos errores que los desarrolladores tienen que corregir.

¿Cómo puede uno asegurarse de que tareas importantes se lleven a cabo a tiempo, y que nadie esté duplicando el trabajo de otro? El sistema de seguimiento de problemas se encarga de esto. Esta herramienta almacena todas las tareas a las que se hace referencia en el sistema, actuando como una especie de lista de tareas. Una tarea puede ser cualquier cosa, desde una solicitud hasta un defecto denunciado por un programador o usuario.

El sistema de seguimiento de problemas permite establecer prioridades, asignar tareas y discutir, con el fin de tomar decisiones informadas sobre cuándo, cómo y por quién, para cada tarea a realizar. La Figura 2 es un ejemplo de problema, o “bug”, sobre el “Like Button” de Facebook. Para cada tarea, un sistema de seguimiento de problema tiene información sobre:

Figura 1

```

Fri Oct 8 10:52:25 2010 Nobuyoshi Nakada <nobu@ruby-lang.org>
    * common.mk (RBCONFIG): depends on version.h due to
      RUBY_PATCHLEVEL. [ruby-core:32709]
Fri Oct 8 00:24:54 2010 James Edward Gray II <jeg2@ruby-lang.org>
    * lib/csv.rb: Fixing documentation typos. [ruby-core:32712]
Thu Oct 7 09:14:28 2010 NARUSE, Yui <naruse@ruby-lang.org>
    * vm_exec.c (vm_exec_core): Treat clang as non gcc on this
      context: It has __asm__ but doesn't works well.
Wed Oct 6 12:28:22 2010 Tanaka Akira <akr@fsij.org>
    * lib/uri/generic.rb (URI::Generic#hostname): new method.
      (URI::Generic#hostname=): ditto.
    * lib/open-uri.rb: use URI#hostname
    * lib/net/http.rb: ditto.
      reported by Adam Majer. [ruby-core:32056]

```

- Un ID de remisión.
- El desarrollador o usuario que lo presenta.
- El desarrollador al que le fue asignada la tarea.
- Una descripción de la tarea.
- Las prioridades, que van de mayor a menor.
- Las gravedades, que van desde triviales hasta críticas.
- Un espacio para las discusiones sobre el tema.
- Y muchos otros espacios con más detalles.

Cuando un commit corrige un problema en el sistema de seguimiento, la práctica común es mencionarlo en el comentario del commit, haciendo referencia al ID en el texto. Esto permite asociar cada error con los cambios reales que lo corrigieron [Fischer2003].

A continuación describimos varios enfoques que se proponen en la literatura que hacen uso de uno o ambos de estos repositorios.

ENFOQUES DE MINERÍA DE REPOSITARIOS DE SOFTWARE

Predicción de cambios

La predicción de cambios responde a la siguiente pregunta: si cambio esta entidad (por ejemplo, una clase o un método), ¿qué otras entidades tengo que cambiar? Si no se cambian estas entidades se puede provocar la introducción de errores en el sistema. La aproximación clásica al problema se basa en el análisis de impacto: uno tiene que explorar todas las entidades que llaman o son llamadas desde un método para determinar si es necesario cambiarlas. El problema es que, además de ser un conjunto potencialmente elevado de entidades a inspeccionar, esto no cubre todos los casos. Por ejemplo, una función que exporta un documento en un archivo no llama a la función de importación, pero por otro lado estos sí están intrínsecamente ligados y necesitan ser modificados conjuntamente.

Figura 2



Una alternativa al análisis de impacto es buscar reglas implícitas en la historia del desarrollo, como se ha almacenado en el sistema de control de versiones. Cuando los desarrolladores suben sus cambios, podemos formar una asociación entre todas las entidades que fueron cambiadas en el commit. Si estas asociaciones se repiten en el tiempo, es posible haber encontrado una regla implícita de programación, como por ejemplo: cuando el método a() cambia, el método b() tiene que cambiar también.

Supongamos que un programador cambia el método a(), pero no b(). Esto vendría a romper el patrón que encontramos. Un enfoque de predicción de cambios puede emitir advertencias cuando los patrones como el de arriba son violados. De hecho, el programador puede saber realmente si se ha olvidado de cambiar b(), lo que constituiría un error que impediría el enfoque. [Zimmerman2004]

Estos resultados pueden ser evaluados con precisión al repetir el desarrollo del sistema. Para cada commit en el sistema de control de versiones, dividimos el conjunto de ficheros que han cambiado en dos conjuntos A y B. Después podemos dar al conjunto del fichero A a uno de los algoritmos de predicción de cambio, y preguntarle cuál es el conjunto B. Como sabemos el contenido de B, podemos compararlo con las predicciones del algoritmo, y medir su performance [Hassan2006].

Predicción de problemas

Mediante el uso de los datos de los sistemas de seguimiento de problemas -es decir, los errores en el pasado- uno puede construir modelos de predicción de errores futuros. El escenario en este caso es el de la asignación de recursos: si un equipo no

tiene suficiente tiempo y/o testadores para verificar correctamente todo el sistema, debe concentrar sus esfuerzos en las partes del sistema que son las más propensas a tener errores. Una vez más, la historia pasada puede ayudar.

Si tenemos un sistema de seguimiento de defectos, sabemos cuántos errores afectan a cada fichero en cualquier momento. Mediante la formación de un modelo de predicción por parte de los datos, y la evaluación de eso sobre el resto de ella, se puede evaluar qué características son mejores predictores de errores en el futuro. Por ejemplo, mientras más grande sea un archivo o más compleja sea una clase, más errores tenderá a tener. Del mismo modo, los archivos que han cambiado más considerablemente en el pasado son más propensos a presentar errores. Hay muchas métricas y enfoques que se pueden usar para esto. Se puede encontrar una comparación de varios enfoques en uno de mis artículos recién publicados [D'Ambros2010].

Recomendación de expertos y triage de errores

Otra área en la que puede ayudar el MSR es la recomendación de expertos. En situaciones de mantenimiento, es común para los desarrolladores tener que realizar cambios en una parte de un sistema que no conoce bien. Sería de gran ayuda tener la opinión de expertos en el área, pero lo que se necesita primero es identificarlos.

Varios indicadores pueden cuantificar la experiencia de alguien, como la cantidad de cambios que realizó a través del tiempo, el número de errores que deberá ser fijado, o el número de veces que utilizó una entidad determinada en su código. Un sistema de recomendación puede ser construido de

modo que, al explorar el código fuente en un IDE, una lista del personal capacitado en el archivo actual se pueda generar con el fin de contactar fácilmente a los expertos, en caso de que sea necesario. La Figura 3, extraída del artículo de Gırba [Gırba2005], muestra cómo los expertos sobre archivos de un sistema cambian con el tiempo y sus actividades. Cada línea representa un archivo y cada punto de color un cambio de un autor. Cada línea tiene el color del autor que sabe más sobre el archivo.

Un problema relacionado es el triage de errores: ¿A quién en el equipo debemos culpar de este nuevo error? Una gran proporción de errores son asignados, efectivamente, a varias personas antes de encontrar a la persona adecuada. Al usar el repositorio defectuoso, podemos entrenar modelos predictivos basados en los errores del pasado. Puesto que sabemos quién en el pasado ha corregido un error, podemos obtener una buena idea del rendimiento real de los modelos de predicción [Anvik2006].

Delta debugging (Delta Depuración o DD) es una forma sistemática para buscar el cambio que causó el accidente basado sobre el sistema de control de versiones. DD realiza una búsqueda dicotómica, mediante la aplicación de la primera mitad de los cambios en el framework, ejecuta la aplicación, y prueba si la aplicación falla. Si es así, refina la búsqueda del error, aplicando el primer cuarto de los cambios. En caso contrario, aplica la segunda mitad de los cambios y hace búsqueda ahí. El proceso continúa sobre el completo historial de los cambios, hasta que DD sea capaz de localizar el commit que es responsable del accidente. Como es evidente, es mucho más fácil entender - y arreglar - un puñado de líneas en lugar de cientos de miles [Zeller1999].

Version-sensitive editing

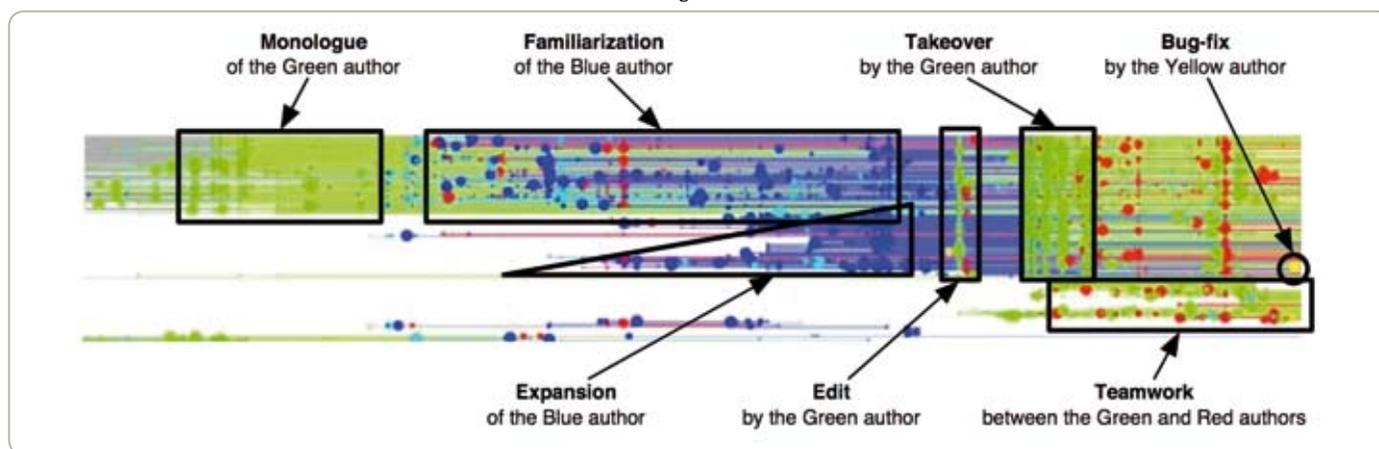
De la misma manera, una de las primeras herramientas que se dedicaba a explorar la información en un repositorio de versiones se

desarrolladores usaron el editor (el editor de versiones firma cada archivo con una firma específica cuando se utiliza), fue posible medir un aumento de la productividad para los usuarios de la herramienta en comparación con los desarrolladores que no hicieron uso de la herramienta. Este incremento, estimado en un 40%, o 1,400 persona/año, se traduce en la cifra mencionada anteriormente [Atkins2002].

MI INVESTIGACIÓN EN MSR

Hasta ahora hemos visto que el MSR puede ser útil en una variedad de casos. Sin embargo, sólo hemos arañado la superficie de todos los enfoques que se han propuesto. Pero hay un principio general: los resultados del MSR son sólo tan buenos como los datos sobre los que se basan. En este contexto, mi objetivo de investigación es hacer frente a este problema, proporcionando datos más precisos, o tipos de datos adicionales.

Figura 3



Delta debugging

Este trabajo fue iniciado por el autor de la herramienta de depuración gráfica DDD que depende del depurador GDB, basado en texto. Entre dos versiones de GDB, DDD dejó de trabajar. Aproximadamente 200.000 líneas de código se cambiaron entre las dos versiones de GDB. La localización de los cambios que son la razón del error es muy difícil en este tipo de situación.

estima que ha salvado 270.000.000 dólares a una gran empresa en el transcurso de varios años. Esta sencilla herramienta, llamada editor de versiones, es un editor de texto que subraya código recientemente retirado, y pone en negritas código recientemente añadido, lo que permite ver más rápido los cambios entre dos versiones.

El editor de la versión se utilizó en una gran empresa durante varios años. Desde que fue posible determinar quiénes de los

Registro de cambios precisos

Los datos almacenados en los repositorios de versiones son notoriamente imprecisos: un sistema de control de versiones trabaja con archivos y no con programas. Si uno quiere hacer un análisis preciso sobre la evolución de los programas, hay que hacer el parsing de cada versión, y después unirlos. Esto representa un montón de trabajo: por

ejemplo, las clases o los métodos pueden cambiar de nombre entre las versiones, lo que es cada vez más común con las herramientas de refactoring en los IDEs. Hay que aceptar la imprecisión o detectarla con un algoritmo específico.

Además, el sistema de control de versiones registró cambios sólo cuando los desarrolladores hicieron un commit; pero ellos quizás hicieron sólo un commit al día, o tal vez uno sólo a la semana. Todos los cambios realizados durante ese tiempo se comprimen en un sólo evento, perdiéndose todas las informaciones sobre el orden de los eventos. Estas dos características de los sistemas de control de versiones reducen en gran medida la precisión de los enfoques MSR que necesitan este nivel de detalle [Robbes2005].

He propuesto una solución al problema de la exactitud de los datos mediante el registro de la actividad de los desarrolladores en el IDE, en lugar de recuperar los cambios de los sistemas de control de versiones. Esto nos permite pensar en las secuencias de cambios en el programa, en lugar de conjuntos desordenados de los cambios en archivos de texto. Esto a su vez nos da una visión mucho más precisa de la evolución del código fuente del sistema. La Figura 4 muestra una visualización de la actividad registrada durante una sesión de desarrollo en la izquierda, y la misma actividad cómo se recupera desde un repositorio de versiones en la derecha. Cada línea representa un método, y cada punto un

cambio a un método (verde: creación de método; naranja: cambio; rojo: supresión; azul: otros). El efecto de “compresión” es evidente.

Doy un ejemplo de uso de esta información: un registro tan detallado de la actividad permite evaluar la eficacia de las herramientas en el IDE, de forma similar pero mucho más precisa que lo que la repetición de desarrollo nos permite (ver la sección sobre predicción de cambios). Esto nos permitió evaluar distintas variantes de herramientas comunes de IDE, como el “completador” de código, y proponer un algoritmo de “completado” mucho más preciso que el algoritmo inicial. De hecho, nuestro algoritmo es casi seis veces más exacto que el inicial. También tiene una interfaz gráfica mejor adaptada a su mayor precisión. Ese algoritmo e interfaz es empleado ahora por todos los desarrolladores que utilizan el entorno de programación Pharo. Hicimos una encuesta sobre las herramientas y los usuarios de Pharo prefieren en gran medida nuestra herramienta (Figura 5, izquierda: herramienta anterior; derecha: nuestra herramienta) [Robbes2010a].

Este trabajo sobre el registro de cambio fue la base de mi tesis de Doctorado, titulada “Of Change and Software” [Robbes2008a]. Esto dio lugar a varias publicaciones en los principales lugares de la ingeniería de software, tales como la ASE [Robbes2008b], ICSE [Robbes2008c] y el Journal de ASE [Robbes2010a], sin olvidar la conferencia MSR [Robbes2007a, Robbes2010b], entre otros.

Hacer un link entre e-mails y código fuente

Además, he investigado la utilidad de otras fuentes de datos tales como archivos de correo electrónico, a fin de tener una visión más completa de la evolución de un sistema. Los archivos de correo electrónico comprenden los debates y las peticiones que los desarrolladores hacen durante toda la vida del proyecto.

El primer paso para explotar estos datos es hacer un link entre los e-mails y los elementos del código fuente al que los e-mails hacen referencia. Uno puede pensar en muchas técnicas -simplicas o complicadas- para hacer esto. Nuestra evaluación de varias técnicas para encontrar links entre e-mails y código fuente se publicó en el ICSE de 2010 [Bacchelli2010] y en WCRE el año anterior [Bacchelli2009] donde obtuvo el premio al mejor artículo. Si el problema parece simple, evaluar sistemáticamente la exactitud de los métodos y sus variantes nos ha permitido encontrar resultados inesperados: enfoques simples basados en expresiones regulares terminó con mejores resultados que las técnicas avanzadas de minería de texto, tales como Vector Space Models y Latent Semantic Indexing.

Nuestro enfoque se centró en e-mails, pero puede funcionar con cualquier repositorio que contiene texto, por ejemplo conversaciones sobre el archivo defectuoso, comentarios de commit, archivos de programas de chat, o sobre la documentación del software.

Figura 4



Figura 5



Figura 6



Minería de ecosistemas de software

Por último, la evolución de un sistema puede verse afectado por otros sistemas, como mostró el ejemplo sobre el Delta Debugging. Actualmente estoy trabajando en minería de la evolución de los ecosistemas de software. Un ecosistema de software es un conjunto de proyectos construidos por una comunidad de software, que evolucionan juntos. Si alguien cambia algo en un sistema - por ejemplo, el nombre de un método - eso puede afectar todos los sistemas que usan ese método. Hemos encontrado casos, incluso en pequeñas comunidades, donde un cambio en un sistema puede tardar hasta seis meses antes de ser adoptado en todos sus sistemas dependientes.

El primer paso en la minería de ecosistemas es recuperar las dependencias entre proyectos. Esto no es fácil, ya que la

cantidad de datos es grande. Nuestra evaluación de técnicas ligeras para recuperar las dependencias entre proyectos, se ha publicado muy recientemente en la ASE de 2010 [Lungu2010]. En la Figura 6, se pueden ver las dependencias que hemos recuperado entre los proyectos de un ecosistema que tiene más de 200 proyectos distintos.

¡Busco estudiantes!

Con el fin de seguir trabajando en estos temas de investigación, busco estudiantes para supervisar. Por lo tanto, si eres estudiante de maestría o de licenciatura interesado en este ámbito y estás dispuesto a invertir parte de tu tiempo en estos temas de investigación, puedes ponerte en contacto conmigo en: rrobbes@dcc.uchile.cl . BITS

REFERENCIAS

- [Lehman1985] M. M. Lehman, L. A. Belady: Program Evolution - Processes of Software Change. Academic Press, London, 1985, pp. 538.
- [Bird2009] C. Bird, P. C. Rigby, E. T. Barr, D. J. Hamilton, D. M. Germán, P. T. Devanbu: The promises and perils of mining git. MSR 2009: 1-10.
- [Fischer2003] M. Fischer, M. Pinzger, H. Gall: Populating a Release History Database from Version Control and Bug Tracking Systems. ICSM 2003: 23-32.
- [Zimmermann2004] T. Zimmermann, P. Weißgerber, S. Diehl, A. Zeller: Mining Version Histories to Guide Software Changes. ICSE 2004: 563-572.
- [Hassan2006] A. E. Hassan, R. C. Holt: Replaying development history to assess the effectiveness of change propagation tools. Empir. Software Eng. 11(3): 335-367 (2006).
- [D'Ambros2010] M. D'Ambros, M. Lanza, R. Robbes: An extensive comparison of bug prediction approaches. MSR 2010: 31-41.
- [Girba2005] T. Girba, A. Kuhn, M. Seeberger, S. Ducasse: How Developers Drive Software Evolution. IWPSE 2005: 113-122.
- [Anvik2006] J. Anvik, L. Hiew, G. C. Murphy: Who should fix this bug? ICSE 2006: 361-370.
- [Zeller1999] A. Zeller: Yesterday, My Program Worked. Today, It Does Not. Why? ESEC / SIGSOFT FSE 1999: 253-267.
- [Atkins2002] D. L. Atkins, T. Ball, T. L. Graves, A. Mockus: Using Version Control Data to Evaluate the Impact of Software Tools: A Case Study of the Version Editor. IEEE Trans. Software Eng. 28(7): 625-637 (2002).
- [Robbes2005] R. Robbes, M. Lanza: Versioning Systems for Evolution Research. IWPSE 2005: 155-164.
- [Robbes2010a] R. Robbes, M. Lanza: Improving code completion with program history. Autom. Software. Eng. 17(2): 181-212 (2010).
- [Robbes2008a] R. Robbes: Of Change and Software. Ph.D. Thesis, University of Lugano, 210 pp.
- [Robbes2008b] R. Robbes, M. Lanza: How Program History Can Improve Code Completion. ASE 2008: 317-326.
- [Robbes2008c] R. Robbes, M. Lanza: SpyWare: a change-aware development toolset. ICSE 2008: 847-850.
- [Robbes2007] R. Robbes: Mining a Change-Based Software Repository. MSR 2007: 15-23.
- [Robbes2010b] R. Robbes, D. Pollet, M. Lanza: Replaying IDE interactions to evaluate and improve change prediction approaches. MSR 2010: 161-170.
- [Bacchelli2010] A. Bacchelli, M. Lanza, R. Robbes: Linking e-mails and source code artifacts. ICSE (1) 2010: 375-384.
- [Bacchelli2009] A. Bacchelli, M. D'Ambros, M. Lanza, R. Robbes: Benchmarking Lightweight Techniques to Link E-Mails and Source Code. WCRE 2009: 205-214.
- [Lungu2010] M. Lungu, R. Robbes, M. Lanza: Recovering inter-project dependencies in software ecosystems. ASE 2010: 309-312.



El DCC entre 1978-1983: Audentes fortuna iuvat*

Presentación en el DCC en 1979. José A. Pino (primero a la izquierda), Francisco Oyarzún y Patricio Poblete (a la derecha).

Fotografía: Gastón Carreño.

En un artículo publicado en el número anterior de esta Revista se relataba la concepción del DCC como Departamento en la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile. Este artículo es una secuela del anterior y recuerda los difíciles primeros años.

El autor regresaba de un posgrado en Estados Unidos y le esperaba hacerse cargo de la dirección del DCC a partir de marzo de 1978. Hasta ese momento, había sido director Fernando Silva, que como se mencionó en el artículo previo, siendo también director del Centro de Computación (CEC) dio pasos cruciales para la fundación del Departamento.

RECURSOS FÍSICOS

A partir de 1973, la Universidad de Chile estaba intervenida. No había Rector, sino rector-delegado, eufemismo con que se designaba al interventor, que excepto muy al comienzo y hacia el final del Gobierno

militar, se trataba de un general de Ejército en servicio activo. En los hechos, los rectores-delegados redujeron la Universidad drásticamente en todos sus aspectos: personal, sedes, presupuesto. Sin embargo, eso no se manifestaba explícitamente. Cada año se "reajustaba" la asignación presupuestaria aportada por el Estado en un monto muy inferior a la inflación del período anterior. Recuerdo un año en que un importante diario local destacaba en un titular: "Aumentan presupuesto de la Universidad de Chile en 28%"; el problema era que la inflación del año precedente había sido 45%. Quizás mucha gente que leyó la noticia no se percató de que el incremento era en términos nominales, no reales, y por lo tanto se trataba más bien de una reducción.

Los rectores-delegados tenían diversas especialidades: infantería, caballería, etc. Hubo uno que era paracaidista. Este último, a poco de ser designado, decidió dictar él mismo un curso libre de paracaidismo.



José A. Pino
Profesor Titular DCC, Universidad de Chile. MS y calificado PhD, The University of Michigan (1977). Co-fundador del DCC U. de Chile. <http://www.dcc.uchile.cl/~jpino>

* "La suerte favorece a los audaces" (Virgilio: Eneida)

Así es que envió una carta a todas las facultades informando del curso, el cual estaba abierto a todos los alumnos de la Universidad. Hasta ahí todo podría parecer normal, pero agregé un párrafo en que especificaba que los alumnos que tomaran dicho curso quedaban eximidos de cualquier evaluación docente en sus carreras que coincidiera con el horario de clases del curso de paracaidismo... Claro, ¡no había comparación de la importancia del paracaidismo con las asignaturas propias de cada carrera! Con todo, los rectores-delegados no se inmiscuyeron en la designación de decanos de nuestra Facultad, excepto un caso a mediados de los '80, pero eso cae fuera del período de análisis.

Con ese telón de fondo, el DCC necesitaba crecer, y mucho. El contraste, en consecuencia, era fuerte. Partíamos con un presupuesto ínfimo, con muy pocos funcionarios, con una planta física muy pequeña, y sin laboratorios de ningún tipo. La creación del Departamento, como una iniciativa “desde abajo”, sin un plan elaborado por personas con poder, significaba que no se destinaron fondos especiales a este proyecto. Fondos que, por lo demás, con las reducciones en curso, en realidad no existían ni en la Facultad ni en la Universidad. A nuestro favor, quizás el recurso disponible más importante del nuevo Departamento era la motivación de sus integrantes.

Con respecto al espacio físico, gracias a las gestiones de Fernando Silva partimos con unas pocas oficinas. Después hicimos crecer este espacio logrando hasta un 25% del primer piso del edificio de Blanco Encalada 2120 (aproximadamente 250 m²). Y sin ningún equipo computacional. El contraste es muy fuerte con la situación actual del DCC, lo que nos debe recordar que a partir de entonces nadie nos ha regalado nada: todo ha sido obtenido con nuestro propio esfuerzo.

LAS PERSONAS

En marzo de 1978 éramos cinco académicos de jornada completa: Miguel Guzmán, Francisco J. Oyarzún, Alfredo Piquer, Patricio Poblete y el suscrito. Todos éramos menores

de 30 años. En el período 1978-1983 se incorporaron los jóvenes académicos Juan Álvarez (permaneciendo desde 1979 a la actualidad), Agustín Araya (1981-1982), José Benguria (1982-1986), Rafael Hernández (1979-1986) y Fernando Taboada (1982-1986). En cuanto a funcionarios, partimos con una secretaria y un auxiliar. Desde el comienzo y por varios años, la secretaria del Departamento fue Arlena Henríquez. Todas las labores presupuestarias, de personal, de Jefe Administrativo, etc., las realizaban funcionarios del Centro de Computación primeramente, y en forma posterior, del Departamento de Matemáticas. Nuestro propio Jefe Administrativo sólo comenzó en 1989: Margarita Serei, quien nos ha acompañado desde entonces. Como teníamos mucha actividad docente, conseguimos un cargo de Secretaria Docente, al cual ingresó Magaly Zúñiga en noviembre de 1981, quien permaneció con nosotros hasta 2009 en ese cargo, jubilándose.

El primer posdoctorado que tuvimos como visitante fue Ernst Leiss (1978). Tiene nacionalidad alemana, pero había obtenido su Doctorado en la Universidad Técnica de Viena, Austria. Venía recomendado por Donald Cowan de la Universidad de Waterloo, Canadá, quien administraba el convenio de apoyo a la Universidad de Chile financiado por IBM. Ciertamente, su estadía de seis meses fue un aporte y un estímulo. En el DCC dictó un curso de “Self-organizing automata”. Posteriormente, viajó a Estados Unidos, y siguió una gran carrera académica en la Universidad de Houston, manteniendo siempre nexos con nosotros y con Latinoamérica en general.

Los sueldos de los académicos eran bajísimos. Definitivamente no alcanzaban para sostener las familias, y quienes la tenían, seguían trabajando para el DCC gracias a un subsidio del cónyuge. En la época se creó una escala mejorada de remuneraciones en la Facultad para los funcionarios del Centro de Estudios Espaciales, con algún argumento de especialización. A esta escala logró adscribirse el CEC, también con el mismo argumento. En la Facultad, se comentaba en broma acerca de los “astronautas del CEC”. Como en alguna conversación le mencioné al Decano el peligro de quedarnos sin

académicos por las bajas remuneraciones, me ofreció, que si queríamos, podíamos también solicitar adscribirnos a esa escala especial (o “espacial”).

Aquí es importante detenernos en este aspecto porque queda explícita la mística del DCC de la época. Queríamos hacer actividad académica, como cualquier otro Departamento de la Facultad. Por lo tanto, argumentamos que después lamentaríamos si entonces pedíamos excepción de lo que realmente era nuestro objetivo. Moralmente, no podíamos ganar más que otros académicos de la Facultad asimilándonos por conveniencia a las remuneraciones más altas de algunos Centros de Servicio. Así es que con pesar, rechazamos la oferta.

El validarnos como académicos frente a la comunidad científica de la Facultad fue uno de nuestros primeros objetivos. Pensamos que una buena idea era tener reuniones con líderes académicos de la Facultad y explicarles lo que queríamos hacer como Departamento científico, y en términos prácticos, diferenciar nuestro quehacer con los servicios de asesoría, uso de computador, mantención, etc. que brindaba el CEC. Sin embargo, las cosas no serían fáciles ni obvias. Nuestro plan inicial se vio frustrado desde la primera reunión. En ésta, después de recibir una breve explicación de lo que queríamos hacer, uno de los académicos senior que habíamos invitado nos respondió: “Les comprendemos completamente. En realidad, no necesitan hacer campaña con nosotros, ya que sabemos perfectamente lo que **Ustedes hacen en el CEC...**” (énfasis del autor). Pasarían años para que el resto de los académicos de la Facultad internalizara efectivamente el hecho de que en Computación se podía hacer investigación científica.

LA DOCENCIA

En estos años estábamos abocados a la docencia en tres niveles. En primer lugar, la docencia básica en el Plan Común de Ingeniería Civil. En particular, siempre dictamos un curso de Introducción a la Computación a los alumnos de Ingeniería. La programación era en Fortran, y

posteriormente en Ratfor (Rational Fortran). Para este curso, creamos apuntes, que puliéndolos, posteriormente se editaron como libro por la Editorial Andrés Bello (J.A. Pino, A. Piquer, P. Poblete: "Introducción a la Computación – Programación Estructurada en Fortran", 1981). Este libro tuvo tres ediciones (Fig. 1).

Figura 1
Texto inicial del curso introductorio



El segundo nivel de docencia era en la Ingeniería de Ejecución en Procesamiento de la Información (IEPI). Este programa de cuatro años fue heredado del Departamento de Matemáticas, cuando se creó el DCC. Al comienzo, había varios programas de Ingeniería de Ejecución en la Facultad, pero de a poco fueron desapareciendo, siendo la IEPI la última activa, terminando su oferta en 1983. Por supuesto, muchos estudiantes se siguieron graduando después de esa fecha, ya que habían ingresado al programa antes. Las Ingenierías de Ejecución fueron implícitamente dejadas a otras universidades que habían comenzado a ofrecerlas, y la Facultad se concentró en las Ingenierías Civiles. La Ingeniería Civil en Computación se comenzó a ofrecer en 1986.

El último nivel de docencia fue el Magíster en Ciencias, mención Computación que, creado en 1975, se ha seguido ofreciendo en el Departamento en forma ininterrumpida hasta hoy.

EL EQUIPAMIENTO

Inicialmente, tanto académicos como alumnos utilizábamos los mainframes IBM del CEC en modo batch. Posteriormente, se ofreció acceso a través de terminales. Sin embargo, su uso era muy restringido considerando estándares actuales. Esta falta de equipamiento fue una de nuestras preocupaciones centrales.

En particular, los alumnos seguían utilizando acceso batch: al comienzo se usaban tarjetas perforadas para los programas, y los resultados se obtenían en papel impreso. La primera mejora (1978) fue el sistema de "Pantallas de Papel", desarrollado por Alfredo Piquer y Patricio Poblete, que permitía que los estudiantes escribieran sus programas en formularios, los que eran digitados por personal del CEC y acumulados en medios magnéticos, para posterior compilación y ejecución en el mainframe. Posteriormente (1980), estos mismos investigadores desarrollaron el sistema SIETE (Sistema Interactivo de Entrada de Trabajos de Estudiantes), que funcionando con terminales, utilizaba pocos recursos del mainframe al atender múltiples usuarios desde una única máquina virtual. Con características similares a Unix, SIETE proveía un sistema jerárquico de archivos.

La falta de recursos económicos era evidente. En 1979 logramos que la empresa Burroughs nos diera "en comodato" un computador B1860 con cinco terminales para uso de los alumnos de especialidad (IEPI y Magíster). El argumento con el que vendimos la idea a Burroughs era de que así los alumnos aprenderían con esa máquina, y muy luego, al graduarse y trabajar en empresas, pedirían el mismo ambiente al que estaban acostumbrados. El valor de lista de los equipos se acercaba a los US\$300.000 comercializado en Chile. Por varios años

fue una solución paliativa de equipamiento, pero al final, se sobrecargaba tanto ese computador que los alumnos obtenían pésimo tiempo de respuesta. Alguien bautizó irónicamente al computador como "la tortuga" y no pasó mucho tiempo para que Burroughs retirara su aporte.

Mientras tanto, habíamos obtenido con mucho esfuerzo algunos fondos de la Facultad para comprar microcomputadores para uso de los alumnos. Importamos directamente para así reducir los costos. Inicialmente compramos nueve computadores Intertec Superbrain (1981), Fig.2.

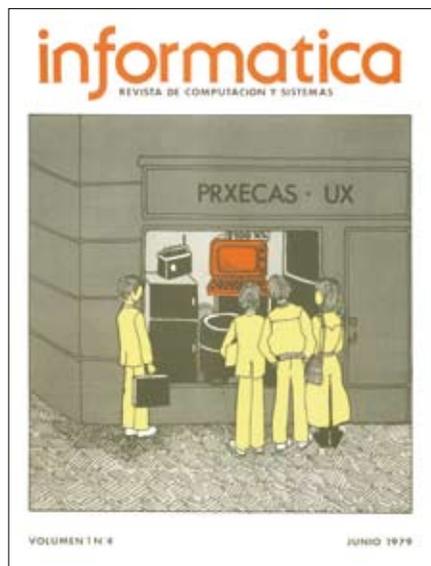
Antes de 1980, los vendedores locales desacreditaban los microcomputadores al calificarlos como "juguetes", hasta que la mismísima IBM sacó en forma sorpresiva ese año su "Personal Computer", iniciando una revolución en el uso de la computación. En la época, el suscrito era además Director de una revista de difusión de circulación local llamada "Informática". Recuerdo que para el número de junio de 1979 imaginé que para algún futuro quizás los computadores se venderían en tiendas, de manera similar a equipos de línea blanca o electrónica de consumo. Le dí la idea al dibujante de la revista y él hizo la portada de ese número (Fig. 3). Después que ese número salió a la circulación, recibí varias llamadas de profesionales del área opinando en el sentido de que claramente estábamos desquiciados: ¡Los computadores jamás se venderían en una tienda!

Figura 2
El microcomputador Superbrain



Foto: Tom Murphy VII (Wikimedia Commons).

Figura 3
Portada de "Informática"



PROYECTOS DE INVESTIGACIÓN

La Universidad de Chile mantenía un Concurso anual de proyectos de investigación. En esos concursos obtuvimos nuestros primeros proyectos. Posteriormente (1986), se creó Fondecyt juntando los fondos de los varios concursos de investigación que había en las universidades del Consejo de Rectores.

Nuestro primer proyecto trascendente en el período fue BIRDS, que tuvo un financiamiento diferente. La historia de este proyecto es como sigue. En 1979, el Gerente de Proyectos de Burroughs Chile era un energético ejecutivo llamado Osvaldo Schaerer. Había un par de grandes clientes que iban a adquirir a esta empresa un mainframe de la línea B6800 y le habían solicitado un software de recuperación de información. El problema era que Burroughs no tenía un sistema para esta necesidad, y los productos en el mercado no ejecutaban en los computadores provistos por esa compañía. Schaerer decidió entonces cotizar en Chile cuánto saldría desarrollar un sistema de este tipo. Preguntó a una casa de software, donde le dieron un presupuesto por hacer un sistema que trataría de replicar la funcionalidad de un software batch conocido en el mercado y que ejecutaba en computadores de la competencia. También averiguó con nosotros en el DCC. Armamos

un pequeño equipo de trabajo compuesto por Alfredo Piquer, Patricio Poblete y el suscrito. Nuestra propuesta era mucho más ambiciosa: tomar el problema original y diseñar una solución interactiva para él, con un lenguaje propio, y servicios avanzados. El software sería además construido en Algol, que era prácticamente el lenguaje de máquina de los computadores B6800, por lo que creíamos que iba a ser eficiente. Se llamaría BIRDS (Bibliographic Information Retrieval and Dissemination System).

A Schaerer le gustó nuestra audaz propuesta, pero en lugar de tratar de financiar el desarrollo con fondos de la sucursal chilena, involucró a la Casa Matriz. Les argumentó que si en Chile los clientes pedían ese software, seguramente a nivel mundial habría un mercado no explotado aún por la compañía. Dentro de poco tiempo, obtuvo la aprobación de la Casa Matriz. Schaerer nos contó la noticia, y con buen humor, nos señaló que el nombre BIRDS estaba bien, pero que la "B" debía ser por "Burroughs". En realidad, una vez completo, el sistema fue comercializado por Burroughs bajo el nombre "TEXT-TRIEVE" (Fig. 4), según nos explicaron porque el nombre "BIRDS" ya estaba registrado en Estados Unidos.

El desarrollo fue relativamente rápido, y con pocas tensiones entre cliente y desarrolladores. Una que recuerdo fue respecto al uso de terminales. Burroughs quería que el sistema funcionara sólo

Figura 4
Uno de los manuales de TEXT-TRIEVE



con los terminales marca Burroughs, que tenían funciones especiales para destacar texto y otras funcionalidades. Nosotros argumentábamos que era mejor que el sistema funcionara con terminales muy básicos, ya que así no sería necesario re-programar parte del sistema cuando se cambiaran los terminales. A regañadientes, aceptaron nuestra solución, pero después Burroughs cambió esta parte para sus terminales específicos. Había una razón comercial para esta decisión, ya que obviamente, los terminales marca Burroughs eran mucho más caros que los terminales genéricos más simples. Sin embargo, después tuvieron que cambiar el software cuando los terminales en cuestión quedaron obsoletos.

Mucho antes que el software estuviera terminado, el incansable Schaerer ya había programado un curso para gerentes de proyectos de Burroughs a nivel mundial en la Casa Matriz de la compañía en Detroit, Michigan, Estados Unidos. Los instructores fueron Óscar Schaerer, Patricio Poblete y el suscrito. Recuerdo una conversación que tuve con un alto ejecutivo de la Casa Matriz en esa ocasión. Estaba interesado en conocer los detalles del desarrollo. Le mencioné entonces que éramos tres desarrolladores. Me replicó que por supuesto, entendía que los tres éramos los ejecutivos del desarrollo, pero que debía haber un grupo grande de profesionales con roles de diseñadores, codificadores, revisores, probadores, documentadores, escritores de manuales, etc. Creyó que era una broma cuando le contesté que no, que éramos sólo los tres. Al término de la reunión quedé con la impresión de que pensaba que no había posibilidad alguna de que a futuro llegara a existir el software del cual estábamos haciendo un curso y que Schaerer le había vendido tan astutamente. Esto era comprensible porque en esa época se acostumbraba en la industria desarrollos estilo "ejército chino", con muchos profesionales involucrados. Y por supuesto, nosotros veníamos de un país sin trayectoria en Computación o en tecnología en general.

Sin embargo, contra estas expectativas negativas, el sistema fue completado a tiempo. Para comienzos de 1981, hicimos

entrega de la versión final de BIRDS, sin perjuicio de que posteriormente hicimos una versión extendida, que fue la memoria de Juan Álvarez. Burroughs siempre quedó muy conforme con el producto. De hecho, nos enviaron una carta diciendo que dentro de su cartera de sistemas de software, TEXT-TRIEVE había sido rankeado como el de mejor mantenibilidad. El sistema efectivamente fue vendido y utilizado en varios países, entre ellos China (más de 20 instalaciones), Australia, Argentina, Bélgica, Estados Unidos y México. Incluso tuve una experiencia gratificante en una conferencia años después. Estaba conversando con un investigador chino cuando, al saber que yo provenía de Chile, me contó que en su centro utilizaban un sistema que les habían dicho había sido desarrollado en nuestro país... ¡era TEXT-TRIEVE!

También años después, nos dimos cuenta de que BIRDS había sido el primer caso de exportación de software de uso masivo que se había hecho en Chile. Burroughs después se fusionó con Univac y cambió su nombre a Unisys. Según me relató Osvaldo Schaerer, trataron de desarrollar ellos mismos una segunda versión de TEXT-TRIEVE en Estados Unidos, recodificando el producto. Sin embargo, nunca pudieron superar o igualar el rendimiento de la primera versión y el proyecto fue finalmente desechado.

LAS CONFERENCIAS

Los primeros eventos de Computación en Chile fueron organizados por la Universidad Católica de Valparaíso en esa ciudad con el nombre "Panel-Expodata". Después, este evento tomaría un carácter itinerante en Latinoamérica con el nombre de Conferencia CLEI.

Para el Panel-Expodata de 1978, los organizadores pidieron que nos involucráramos en la Conferencia, lo que hicimos. Especialmente motivado estaba Francisco J. Oyarzún, quien tenía un proyecto llamado BIOS, de simulación de seres vivos. Con sus ayudantes de investigación envió una gran cantidad de artículos en respuesta al llamado a Presentación de Trabajos, lo cual descolocó a los organizadores. La solución de organización de sesiones que encontraron

fue la siguiente. Los artículos aceptados se organizaron en tres grandes bloques paralelos: ¡Hardware, Software, y BIOS! Imagino que hasta ahora esta clasificación debe ser única en el mundo.

Considerando que para 1979 el Panel-Expodata iría a otro país, pensé que se podría organizar un evento en Santiago en ese año. Conversé con personas del Centro de Ciencias de la Computación (CECICO) de la Universidad Católica y aceptaron co-organizar con nosotros la Primera Conferencia Nacional en Teoría de la Computación y Desarrollo de Software. Se efectuó en agosto de 1979 y fue bastante exitosa en términos de artículos y asistentes, por lo que decidimos repetir la fórmula el año siguiente (1980). Sin embargo, se prefirió cambiar el nombre y así se celebró la Segunda Conferencia Nacional en Sistemas de Computación. A este evento vinieron Gerald Estrin, Dan Berry y Carlos Lucena.

En los dos años siguientes, fuimos más ambiciosos y co-organizamos respectivamente la Primera y Segunda Conferencia Internacional en Ciencia de la Computación. En todos estos casos la conferencia fue en Santiago, usando dependencias de nuestra Universidad o de la Universidad Católica. Para la conferencia de 1982 vinieron Dan Berry, Ernst Leiss, Erol Gelenbe y Larry Kerschberg.

Estos eventos y la colaboración para la organización fueron precursores de las conferencias de la Sociedad Chilena de Ciencia de la Computación (SCCC), la que se constituyó con posterioridad.

EL DÍA EN QUE EL DCC DESAPARECÍA

Un día cualquiera de 1981 recibí una llamada de parte del Decano, pidiendo si en lo posible, podía ir de inmediato a su oficina. Cuando llegué, el Decano, Claudio Anguita, y el Vicedecano, Augusto León, me esperaban con cara de funeral. Sin mayores preámbulos, me pasaron una carta. Era del rector-delegado. En ella decía que había un pequeño grupo de oficiales de Ejército que había seguido un curso de Administración en la Academia de Guerra.

En dicho curso, habían aprendido que una buena administración implicaba que todo jefe debería tener siete subordinados, con una tolerancia de -2 o +2. Los oficiales en cuestión habían hecho su práctica en la Universidad de Chile y le habían informado al rector-delegado que habían facultades en que los decanos tenían un exceso (es decir, más de nueve) subordinados (directores de Departamento). La carta finalizaba diciendo que, dado que la Facultad de Ciencias Físicas y Matemáticas era una de estas facultades infractoras, su Decano debía tomar las acciones para corregir la situación.

A mí me pareció un chiste de mal gusto y así lo dije. Es verdad que en Administración se habla del número ideal siete, pero en ningún caso es para tomarlo literal. Sin embargo, el Decano me advirtió que esto era muy serio. Él había llamado al Decano de Medicina y éste le había señalado que respondería al rector-delegado que lo pedido era imposible, puesto que la Facultad de Medicina tenía 65 departamentos. Pero nuestro Decano no quería provocar una ruptura.

La estrategia de la Decanatura era hacer un gesto al rector-delegado, aunque no acogiendo totalmente la petición. La idea era reducir en tres los 17 departamentos existentes. Se juntarían los departamentos de Geología y Geofísica (no tenían mucho que ver, pero empezaban por "Geo"). También los Departamentos de Ingeniería Química y Química Básica ("igual son químicos"), y aquí venía la bomba, los Departamentos de Ciencias de la Computación y Matemáticas. Más aún, el Decano y Vicedecano habían pensado que yo era la persona ideal para dirigir este último departamento fusionado.

Mi reacción inmediata fue de rechazo. Expliqué que para nosotros, el tener un Departamento diferenciado era una cosa fundamental: era un hecho destacable que en nuestra Universidad se reconociera pioneramente la identidad de la disciplina. ¿Por qué no fusionaba los Departamentos de Física y Astronomía? El Vicedecano me reconvino para que no me pasara de listo (el Decano era astrónomo). El Decano entonces me explicó que el Departamento de Astronomía tenía muchos convenios internacionales que lo mencionaban

explícitamente en los contratos. Y que meditara lo que habíamos hablado con más calma.

Esa tarde tuvimos una larga asamblea de académicos de jornada completa del DCC. Todos estábamos de acuerdo en que la solución propuesta por la Decanatura era inaceptable. ¿Significaba que renunciábamos todos y cerrábamos la cortina del DCC? Aunque era duro reconocerlo, parecía ser la única salida coherente y digna. Esto significaba que debíamos empezar a buscar un nuevo trabajo, con mucho dolor porque queríamos la Universidad y más aún nuestra creatura: el DCC. ¡Cuántos sueños quedarían trancos y cuántos estudiantes quedarían a medio camino porque alguien había interpretado mal una idea trivial en un libro de Administración! Esa noche, creímos que el DCC moría.

Al día siguiente pedí reunirme con el Decano. Le conté nuestra decisión. Me dijo que no lo tomáramos así, que algo seguramente se podría hacer. Que no nos apresuráramos, y que buscáramos una solución intermedia aceptable.

Al otro día me llamó, porque tenía una nueva proposición. Consistía en lo siguiente. La Facultad informaría al rector-delegado la fusión de los tres pares de departamentos. Sin embargo, internamente en la Facultad, cada parte de un departamento fusionado funcionaría por separado: tendría su Director, su coordinador docente y su presupuesto. Los directores adicionales serían miembros del Consejo de Facultad con plenos derechos, porque el Decano tenía la atribución de nombrar unos pocos miembros a su arbitrio. Seguiríamos en el fondo, igual, pero a Casa Central le presentaríamos esta “ficción” del departamento fusionado. A cada parte se le llamaría “División”, término que no existía en la estructura oficial de la Universidad, pero dentro de la Facultad todos sabríamos de qué se trataba.

Volví a la asamblea de académicos. Nuestro consenso fue que no obtendríamos mejor posición en una negociación. Así es que aceptamos. Desaparecía el Departamento, pero nacía la División Ciencias de la Computación (DCC), al menos por un tiempo.

El Decano cumplió su palabra y afortunadamente no tuvimos grandes dificultades, excepto en la breve Decanatura de Juan Antonio Poblete (1985), quien no fue aceptado por la comunidad de la Facultad. El Departamento refundido se llamó Departamento de Matemáticas y Ciencias de la Computación, y su director (siempre un matemático) nunca creó la “División Matemáticas”, que hubiera correspondido. Nosotros intencionalmente ignorábamos la dependencia oficial a este “Departamento” y nuestro membrete de correspondencia siempre hacía exclusiva referencia a la “División Ciencias de la Computación” (Fig. 5). Todo el entuerto recién se reparó con el retorno a la Democracia, en que volvimos a tener oficialmente el añorado Departamento de Ciencias de la Computación según decreto de Rectoría del 25 de octubre de 1989. El Departamento de Matemáticas se llamó a partir de entonces Departamento de Ingeniería Matemática.

EPÍLOGO

Varios de los académicos de jornada completa que habían ingresado al DCC, en algún momento decidieron seguir sus carreras en otras partes. Diversas razones existían para ello. Además de los bajos sueldos, ya mencionado, concurrían aspectos vocacionales, oportunidades de trabajo profesional excelentes, emigración hacia otros países, etc. En el caso de Francisco J. Oyarzún se dio la situación que decidió ir a estudiar un doctorado en Biología en Estados Unidos. El DCC le negó el apoyo para mantenerle el sueldo, ya que con nuestro exiguo presupuesto queríamos

Figura 5
Rótulo de correspondencia oficial de la División Ciencias de la Computación



financiar académicos de Computación, no de otras disciplinas, aunque fuesen muy importantes. La Decanatura, sin embargo, le otorgó la mantención de sueldo de su propio presupuesto. El académico no volvió a la Facultad. Otros jóvenes fueron después enviados a seguir programas de doctorado en Computación en el exterior para que regresaran como académicos, y así lo hizo la gran mayoría de ellos.

Entre los costos, el suscrito sufrió el estrés de dirigir el Departamento que contaba con tan pocos recursos, pero con tanta audacia y tantas expectativas. A fines de 1982 se le diagnosticó una úlcera gástrica y la recomendación de cambiar de actividad. Elegimos entonces un nuevo Director, que resultó ser Jorge Olivos, un matemático brillante que regresaba de hacer su doctorado en Computación en Francia, y quien asumió con mucho entusiasmo a comienzos de 1983.

Así se cerraba el período 1978-1983, que había sido complicado, pero que no había disminuido nuestra motivación. La determinación del puñado de académicos, la respuesta siempre muy positiva de los alumnos además de un poco de suerte habían logrado que el DCC se estableciera y progresara. Más aún, el DCC se preparaba para etapas de crecimiento más fuerte en mejores épocas futuras de Chile. BITS



En la foto actual, de izquierda a derecha: Patricio Poblete, José A. Pino, Margarita Serei y Juan Álvarez.

En 20 años de computación: una visión muy personal

Leopoldo Bertossi (a la izquierda) durante la inauguración de las 2^{as} Jornadas Chilenas de Computación (1994).



Leopoldo Bertossi

*Profesor Titular de la Escuela de Ciencia de Computación de la Carleton University (Ottawa, Canadá).
Faculty Fellow del IBM Center for Advanced Studies y miembro del ACM Distinguished Speakers Program.
Profesor jornada parcial en el Programa de Posgrado en Computación de la Universidad de Concepción.
bertossi@scs.carleton.ca*

No es un tema simple recibir una invitación a escribir un artículo¹ que mezcle un poco de la trayectoria personal con la del área a la cual uno ha dedicado muchos años de trabajo, energía, entusiasmo y vocación. No es fácil porque, primero, se refrescan muchas cosas en la memoria y en los recuerdos, de lo placentero y de lo menos grato. La segunda razón es que no es precisamente a la gente joven a la que le piden que escriba este tipo de reseñas. Otro tipo de evidencia, entre muchas otras, del proceso de envejecimiento en el que nos encontramos.

Mi vida académica, y algo más que eso también, ha estado sumergida en la actividad científica; sobre todo, en la matemática y la Ciencia de la Computación. Funciono y escribo desde la perspectiva de un

matemático que ha intentado, primero, comprender ciertas áreas de la Ciencia de la Computación, y, luego, contribuir al desarrollo de algunas de ellas. El impulso de comprender ha estado siempre primero. El intentar esa comprensión, ese darle sentido, significado y estructura a los conceptos, y el ordenarlos mentalmente lleva también a desarrollarlos y a contribuir con nuevas ideas.

Hace prácticamente veinte años hice la transición desde la matemática propiamente tal hacia la Ciencia de la Computación. Este es un buen momento para reflexionar sobre lo que he encontrado en mi nueva área y en la comunidad que la practica. Pareciera que veinte años no es nada en una trayectoria académica. Sin embargo, la computación como disciplina académica es aún bastante

¹ Este artículo surgió de una invitación a escribir una columna o artículo corto sobre algo de mi trayectoria y visión de la comunidad académica de computación. Es claro que la extensión final no era la esperada por el medio de publicación. Sin embargo, agradezco el impulso recibido. He incluido en este artículo varias cosas que considero importantes y lo suficientemente relacionadas como para no querer omitirlas.

joven en Chile; y no hay muchos en el país con una trayectoria de veinte años o más que involucre de manera sostenida la investigación en computación.

EN UN COMIENZO

Para explicar e iluminar la perspectiva personal desde la cual hago mis observaciones y apreciaciones, debería entregar algunos antecedentes sobre mi trayectoria académica. El primero de ellos es que realicé todos mis estudios primarios y secundarios en el Liceo Experimental Manuel de Salas (entonces dependiente de la Universidad de Chile). Esto fue bueno y malo, pero nunca irrelevante ni sin consecuencias ni carente de interés. Entre lo bueno, destaco el espíritu crítico que se nos inculcó, a veces, hasta iconoclasta. Ahí era perfectamente aceptable el salirse de la corriente; aun el ir en contra de ella. También se enfatizaba la autodisciplina. En muchas materias impartidas, el nivel era excelente, a nivel académico en contenido y forma.

A fines de la enseñanza media (¿o era todavía humanidades?), habiendo pasado por la especialización en “matemática-física”, ya tenía claro que me interesaba la ciencia, especialmente la química. En matemática me iba muy bien, era capaz de aplicar bien toda la operatoria, de resolver problemas, incluso de demostrar teoremas, pero sentía que no la entendía y que me costaba más. En particular, las demostraciones geométricas me parecían particularmente misteriosas, ya que nunca estaba claro qué se podía suponer y usar.

Decidí entrar a la Pontificia Universidad Católica de Chile (PUC) a estudiar ingeniería, pensando en especializarme en química. Me di cuenta de dos cosas a poco andar. La primera, que no encontraría mucha ciencia química en la carrera de ingeniería química. La segunda, que la matemática era algo distinto a lo que había aprendido en el liceo. Que detrás de las operaciones algebraicas había unos pocos axiomas o principios que las explicaban todas. También, que las

demostraciones matemáticas seguían ciertas leyes y tenían un origen muy claro. Se me abrió todo un mundo fascinante. Además, mi tendencia natural hacia la abstracción calzaba perfectamente con la matemática, en términos de temas y de tipo de actividad. Hacia el segundo año de universidad, y habiéndome ido muy bien en los estudios de ingeniería, me cambié internamente a estudiar licenciatura en matemática. Me fue muy bien en los cursos, pero no puedo decir que me era fácil. Sin embargo, el desafío me resultaba estimulante, y los contenidos, altamente interesantes.

Durante mis estudios de licenciatura las áreas que se me hicieron más difíciles fueron la lógica matemática y las probabilidades. A propósito no digo “teoría de probabilidades” que, como teoría matemática, no es necesariamente más compleja que el resto. Me refiero también al sentido y al uso del concepto de probabilidad. Eso sí que me resultaba más complejo y misterioso. Curiosamente, tanto la lógica matemática como las probabilidades, tienen, en mi opinión, un alto contenido filosófico. Durante la licenciatura me especialicé informalmente en probabilidad y estadística matemática. Y durante mis estudios de posgrado, terminé trabajando en probabilidad y en lógica matemática. Estas áreas siguen estando cerca de mi corazón. Pareciera ser que mis elecciones se van siempre por el lado de lo que me es más difícil y desafiante.

En el verano de 1975 hice una práctica de vacaciones (voluntaria, ya que no era requisito para la licenciatura en matemática) en la Empresa Nacional de Computación (ECOM). En esa época la computación en Chile estaba fuertemente concentrada en esa empresa estatal. Como no tenía conocimiento alguno de computación, ni siquiera de programación básica, fui asignado al Área de Estadística de la División de Capacitación y Asistencia Técnica, que se dedicaba a la capacitación de usuarios internos y externos, y a prestar servicios externos. Ahí me sentí muy a gusto, ya que había un ambiente estimulante y académico,

incluyendo una muy buena biblioteca. Al final del verano me ofrecieron quedarme tiempo completo como consultor.

Mi trabajo en ECOM consistía en aprender por mi cuenta técnicas estadísticas avanzadas, y usarlas por medio de paquetes computacionales estadísticos, como SPSS, enseñarlas a otros colegas; y en escribir manuales de uso, incluyendo los conceptos y técnicas básicas. También tenía que usar esos paquetes en el procesamiento de encuestas estadísticas para usuarios externos e interactuar con estos últimos. Se me permitía asistir a mis clases de licenciatura en la universidad, pero combinar las dos cosas se me hizo muy pesado.

Hacia el final del año el gerente de la división me forzó a inscribirme en el Plan Nacional de Capacitación en Informática, impartido por la misma ECOM, para que me convirtiera en “analista de sistemas”, una salida profesional práctica muy común en esa época. Tuve muy claro que esa línea era incompatible con la trayectoria que me había trazado en matemática, y, en consecuencia, renuncié a ECOM. Afortunadamente en ese momento, a un semestre de terminar mi licenciatura, fui contratado por el Instituto de Matemáticas de la PUC, como profesor de media jornada.

Al comienzo de mis estudios de posgrado consideré el especializarme en física-matemática, lo que no es extraño. Siempre he tenido mucho respeto y aprecio por la física. Además, siempre me he sentido cercano a lo que se llama usualmente “matemática aplicada”. Sin embargo, al mismo tiempo comencé a leer por mi cuenta sobre lógica matemática, volviendo, con otras fuentes, sobre los temas que me habían resultado antes tan difíciles. Y decidí entonces dedicarme a esa área. Mi tesis de magíster fue sobre fundamentos lógicos de la probabilidad con aplicación al modelamiento del movimiento Browniano. Obtenido el Magíster, mi contrato fue extendido a tiempo completo, en lo que había pasado a ser entretanto la Facultad de Matemáticas de la PUC.

Ya inmerso en la lógica, y también en forma autodidacta, me interesé en la conexión entre lógica matemática y computación. Hay que enfatizar que los orígenes de la computación como disciplina científica están en el trabajo de lógicos matemáticos como Alan Turing, Kurt Gödel, Alonso Church, Emil Post, Stephen Kleene, etc., de principios de los años '30. No hay que olvidar la motivación entregada a estos investigadores por David Hilbert en las tres primeras décadas del siglo veinte, con sus trabajos sobre lógica, fundamentos de la matemática, y problemas algorítmicos. A través de los estudios autodidactas que hice durante mi Doctorado se despertó mi interés en la Ciencia de la Computación.

Durante todos mis estudios de Doctorado seguí leyendo sobre los temas de computabilidad, complejidad computacional, y programación en lógica, a pesar de que no eran parte de mi investigación. Mi tema de tesis era sobre teoría de modelos topológicos, un área de la lógica matemática. Los modelos estudiados en la lógica clásica se pueden ver como generalizaciones o abstracciones de las estructuras algebraicas. En mi caso, los modelos eran generalizaciones de estructuras topológicas; y mi investigación giraba en torno a las propiedades expresables en lenguajes topológicos que se preservan bajo transformaciones topológicas. A comienzos de 1988 terminé mi Doctorado, con la decisión de dedicarme en el futuro inmediato a la Ciencia de la Computación.

EN TRANSICIÓN

A partir de ese momento y durante 1988 las cosas se desarrollaron de manera bastante vertiginosa. Junto con Javier Pinto, y por iniciativa de él, iniciamos un taller de lectura de papers claves en representación lógica de conocimiento. Esta área tiene que ver con la representación de conocimiento en el computador y con su uso en sistemas computacionales. Javier era en ese momento un profesor joven en el Departamento de Ciencia de Computación (DCC) de la Escuela de Ingeniería de la PUC; y había regresado hace poco de sus estudios de máster en la Universidad de California en



Con Javier Pinto en Lisboa.

Los Ángeles. Esporádicamente hubo otros pocos participantes, pero la mayor parte del tiempo y de manera regular nos juntábamos solamente los dos. También dictamos de manera conjunta un curso de posgrado en el tema para alumnos del DCC.

Esta fue una etapa provechosa, donde aprendí cosas nuevas que me resultaron profundamente interesantes. Sobre la base de esas reuniones escribí mi primer paper en Ciencia de la Computación, más precisamente, sobre representación de conocimiento.

Ingresé a la Sociedad Chilena de Ciencia de Computación (SCCC) y participé en la Conferencia Internacional de la SCCC. Especialmente memorable y motivador fue un tutorial sobre "lógica y bases de datos" dado por Alberto Mendelzon, profesor de la Universidad de Toronto.

Conseguí fondos de la Fundación Andes para financiar una estadía larga de investigación en la PUC del profesor Joachim von Zur Gathen, también de la Universidad de Toronto, quien quería pasar parte de su año sabático en Chile. El dictó un curso muy interesante sobre complejidad computacional algebraica.

Fue precisamente a través de Von Zur Gathen, después de su estadía en Chile, que conseguí un puesto de profesor asistente visitante, por un año, en el Departament

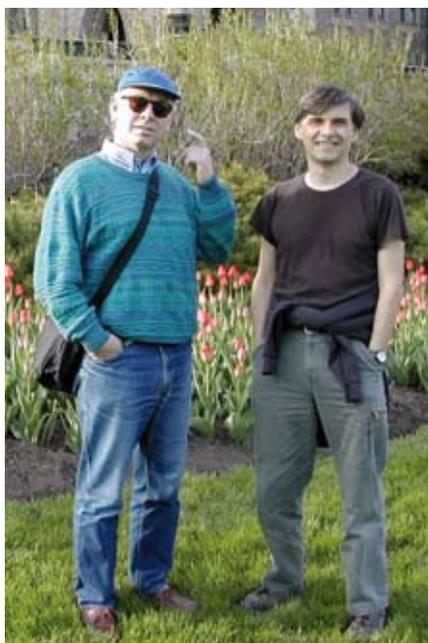


Con Javier Pinto (a la derecha).

of Computer Science de la Universidad de Toronto. Éste fue más bien un salto mortal al vacío si se considera la pobreza de mis conocimientos de computación en esos momentos. Decidí asumir el riesgo y la dureza segura de lo que sería una experiencia de ese tipo. Sentí que esa era la oportunidad inmediata que se me abría para hacer la transición a la Ciencia de la Computación. Y sobre esto último no tenía duda alguna.

Solicité permiso sin sueldo a la Facultad de Matemática y, con el estómago apretado, partí a Toronto en agosto de 1989. Mi esperanza era poder hacer investigación en Toronto, además de mis tareas docentes. El mismo día Javier Pinto viajó a Toronto a iniciar sus estudios de Doctorado en Ciencia de la Computación.

Por supuesto, la presión a la que estuve sometido en Toronto fue peor de lo que había imaginado. A poco de haber llegado empecé a dictar un típico curso de primer año de introducción a la computación. Este incluía programación en el lenguaje Turing. Junto con dictar las clases tuve que aprender todo desde cero, y sólo un poco antes que los alumnos. Simultáneamente dicté el curso de pregrado, muy interesante y atípico, de matemática discreta para computación. El que fuera de matemática no me hacía la tarea más liviana, ya que la mayoría de los temas no los había visto nunca en mi



Con René Peralta (a la derecha) en Ottawa.

vida. Durante el segundo semestre de mi estadía dicté un curso sobre algoritmos, complejidad y computabilidad. Al menos los dos últimos temas no me eran desconocidos. Sin embargo, la experiencia docente fue, en realidad, horrible (afortunadamente más para mí que para los alumnos).

Ahora, por el lado más positivo, aprendí muchas cosas interesantes, útiles, y básicas de computación. Además, durante ese año me hice el tiempo para asistir a varios cursos y seminarios del Departamento, y a las conferencias de los múltiples investigadores que pasaban de visita. El ambiente académico era estimulante y del más alto nivel.

Como será fácil imaginar, no tuve mucho tiempo para hacer investigación. Además, yo era visita oficial del grupo de teoría de la computación, pero mis intereses me llevaban más bien hacia las aplicaciones de la lógica en inteligencia artificial y bases de datos. Esto me tenía en una situación incómoda. Afortunadamente, hacia el final del año en Toronto, Javier Pinto me ofreció ponerme en contacto con su supervisor, Ray Reiter. Hay que destacar que Reiter era figura protagónica a nivel mundial en inteligencia artificial, y representación de conocimiento, en particular.

Con Javier habíamos estudiado con fascinación y cuidadosamente algunos de sus trabajos. Los tres fuimos a almorzar, un día viernes, recuerdo. Ahí Reiter me dijo que tal vez podría ayudarlo con mi formación en lógica matemática, ya que tenía un problema lógico -un presunto teorema- que él creía que era cierto, pero que no podía demostrar. Fuimos a su oficina, me formuló el problema de manera muy simple. Era la oportunidad que había estado esperando y no podía desaprovecharla. Desde ese momento y el fin de semana completo me concentré en el problema y logré resolverlo. El lunes siguiente, después de ver la solución, Reiter me invitó a que trabajáramos en otro tema. Así se inicio mi colaboración con uno de los investigadores que más he admirado. Fue muy motivador, inspirador y formativo el trabajar con él. En ese momento no estaba muy consciente de que trabajar con Reiter me abría puertas y oportunidades valiosas.

La lógica clásica, aquella en la que se basa el razonamiento deductivo matemático, por ejemplo, tiene la propiedad de monotonía. Esto quiere decir que, si a un conjunto de axiomas o a una base de conocimiento se le agrega nuevos axiomas, los teoremas o las consecuencias lógicas que se obtenían de la original siguen siendo válidos con respecto a la teoría expandida. Sin embargo, el razonamiento con sentido común que realizan los seres humanos es esencialmente no monótono. El ejemplo clásico es el de concluir o conjeturar, a partir de la información de que Piolín es un pájaro, que éste vuela. Ésta es una conclusión posiblemente provisional, que puede ser necesario descartar si se agrega la información de que, por ejemplo, Piolín es un avestruz. Los seres humanos necesitamos este tipo de razonamiento para funcionar en un ambiente que ofrece información incompleta. La modelación lógica del razonamiento con sentido común y su adopción en sistemas computacionales que exhiban inteligencia (artificial) ha sido y sigue siendo un gran desafío. Ray Reiter hizo importantes contribuciones en el área

de lógicas no monótonas para representación de conocimiento de sentido común.

Con Reiter trabajé en la aplicación de circunscripción, una forma de razonamiento lógico de sentido común, a la modelación y procesamiento computacional de la noción de objeto genérico. En especial nos interesó el uso de este concepto en el contexto de una teoría matemática, concretamente la geometría clásica. Este tipo de investigación fue del tipo que más me atrae y acomoda. Primero, hubo que intentar entender y estudiar una noción que se maneja intuitivamente. En este caso, la de objeto genérico. Piensen, por ejemplo, en un árbol genérico, o en una casa genérica, de esas que dibujan los niños. O en un triángulo genérico, de esos que se dibujan de manera auxiliar para conjeturar un teorema o guiar una demostración. En segundo lugar, fue necesario formular (definir, conceptualizar, caracterizar, especificar, ...) el concepto en términos lógicos. Una vez hecho eso, restó demostrar que el modelo entrega los resultados esperados, es decir, que tiene las propiedades intuitivamente esperadas. Esto requiere de un análisis matemático.

Dentro de este trabajo, también establecimos que había conexiones interesantes con otros temas aparentemente desconexos, en nuestro caso, con ciertas formas superficialmente paradójicas de razonamiento probabilístico. El establecer puentes entre áreas no conectadas de manera obvia es algo que siempre me ha atraído. Escribí dos papers con Reiter, aprendí mucho trabajando con él, y se generó una amistad entre ambos que duró hasta el momento de su prematura muerte el año 2002.

En lugar de volver a la PUC después del año en Toronto, solicité una extensión del permiso sin sueldo, y me fui como profesor visitante por un semestre al Department of Computer Science de la Universidad de Wisconsin en Milwaukee. Esa oportunidad me la dio mi amigo René Peralta, quien era profesor en ese Departamento. Dicté un curso de inteligencia artificial, otro de matemática discreta para computación,

y supervisé a un alumno de máster. En investigación me concentré en el trabajo iniciado con Reiter.

A principios de 1991 volvía a la PUC, reintegrándome a la Facultad de Matemática (FM). Sin embargo, tenía claro que mi permanencia en ella sería pasajera. A pesar de esto, intenté revitalizar la Ciencia de la Computación en la FM. Me reintegré al Programa de Computación de la FM, que era un resabio pequeño y sólo parcialmente formal de lo que había sido el Departamento de Ciencia de Computación de la FM, el que había desaparecido en 1983 (ver la sección siguiente). Dentro de lo posible en el contexto de la programación docente de la FM dicté cursos relacionados con Ciencia de la Computación. De hecho, la mayor parte de mis actividades las concentré de manera informal en el Departamento de Ciencia de Computación (DCC) de la Escuela de Ingeniería. Dicté el curso de “Lógica para Ciencia de Computación”, que era obligatorio para los alumnos de Ingeniería Industrial con mención Computación. Fui el primero en dictar ese curso, que había sido creado antes de mi regreso a Chile en 1991. Hice de editor del número dedicado a la computación de la revista de la Escuela de Ingeniería, dicté un curso de bases de datos para el postítulo del DCC, tomé contacto con alumnos del departamento, etc. En la práctica, funcioné como un profesor más del DCC, pero sin contrato.



Inaugurando las 2^{das} Jornadas de Chilenas de Computación (1994).

Paralelamente, presenté a la FM un plan de desarrollo de la Ciencia de la Computación. Sin embargo, éste no tuvo acogida, lo cual era esperable y, como veremos, también comprensible. Decidí buscar otros horizontes. Postulé a un concurso abierto por el Departamento de Informática de la Universidad Técnica Federico Santa María (UTFSM). Entremedio, el DCC de la Escuela de Ingeniería de la PUC abrió una vacante, a la cual postulé. El proceso fue muy rápido y se me ofreció el puesto. Sin embargo, a pesar de llevar años en la PUC, se me tomó a prueba por dos años, y se me hizo pasar por un examen psicológico, como a todos los nuevos profesores de la Escuela de Ingeniería. Como ya había decidido irme de la FM de todos modos, acepté esas condiciones. No mucho después la UTFSM me ofreció el puesto al que había postulado, pero ya había aceptado la oferta de la PUC. Dejé la FM y me integré al DCC en 1992.

EN UN DEPARTAMENTO DE COMPUTACIÓN

Hasta 1981 toda la escasa actividad académica en la PUC en torno a computación estaba concentrada en la división de servicios computacionales de la Universidad, que no era una unidad académica. Se impartían los cursos de programación para distintas carreras, generalmente por alumnos de ingeniería y licenciaturas científicas que se habían especializado en la práctica en esos aspectos de la computación. La docencia no era precisamente su fuerte; y la complejidad administrativa de tomar uno de esos cursos no los hacía precisamente atractivos. Había que lidiar con formularios de programación, con cajas de tarjetas, y listados eternos y crípticos que mostraban la corrida del programa. En estos últimos había que sumergirse en la búsqueda de los errores que habían impedido que el programa corriera en la forma esperada. Está bien que un curso sea difícil porque el contenido temático lo es, pero no que las fuentes de complejidad vengan por otros lados.

En 1981, Rolando Chuaqui, Decano de la Facultad de Matemáticas (FM) de la PUC, impulsó la creación del Departamento de Ciencia de la Computación (DCC), como uno de los tres departamentos de la FM. Este fue el inicio de la actividad académica institucionalizada en computación en la PUC. Esta iniciativa, no carente de sobresaltos, fue apoyada tíbiamente por los profesores establecidos de la FM. Había temas complicados, como por ejemplo, la naturaleza no siempre científica de la computación. Otro era la calificación y vocación académica de los profesores que se contrataron en forma ad hoc. La mayoría de ellos eran ingenieros que habían exhibido interés en la computación, y más en su parte práctica y técnica (principalmente programación) que científica. La idea era que ellos salieran al exterior en el corto plazo a obtener un Doctorado en Ciencia de la Computación. Otro tema que produjo controversias fue el de los sueldos, ya que la gente de computación aspiraba a sueldos más altos que el promedio de la FM. Se creó la carrera de Licenciatura en Matemática con mención Computación.

No mucho después, la Escuela de Ingeniería (EI) de la PUC comenzó a interesarse en el desarrollo de la computación como una de sus disciplinas. Hubo conversaciones y acercamientos entre la EI y los profesores del DCC de la FM. Al final, se dieron las condiciones para que la EI propusiera a la rectoría de la PUC que el Departamento de Ciencia de la Computación, y toda la actividad académica en computación, se concentrara en la EI. En 1983 la FM perdió el DCC por decreto de rectoría. Los profesores de la FM asociados al DCC pudieron elegir entre quedarse en la FM o irse al nuevo DCC en la EI. Sólo uno de los profesores, René Peralta, quien estaba haciendo su Doctorado en Computación en Berkeley, decidió quedarse en la FM. De nada sirvieron las objeciones del Decano Rolando Chuaqui. A lo más pudo dejar constancia por escrito, en el documento base para el decreto de rectoría, de sus aprensiones con respecto al futuro poco



Leopoldo Bertossi, Ray Reiter y Jorge Lobo, 1^{ras} Jornadas Chilenas de Computación (1993).

promisorio que veía para la Ciencia de la Computación, como disciplina científica, si ésta se radicaba solamente en la EI.

Estos eventos en torno a la computación fueron bastante traumáticos para la FM. En ella se formó una figura más bien informal de Programa en Computación, muy pequeño, y más bien para dar soporte a la licenciatura que se había creado (aunque la docencia en computación para sus alumnos quedó a cargo del DCC de la EI).

El DCC de la EI se formó eclécticamente, y más de hecho que por diseño, con los profesores que emigraron de la FM, más algunos ingenieros jóvenes que trabajaban en la división de servicios computacionales y en la misma EI en forma más bien aislada, proveyendo apoyo computacional. Sólo un profesor del nuevo departamento contaba con un Doctorado en Ciencia de la Computación, de hecho, muy reciente. Los otros no tenían Doctorado, y, tal como era el plan en la FM, ellos saldrían a hacerlo en un corto plazo.

Vale la pena hacer notar que la EI acostumbraba contratar ingenieros jóvenes -sus alumnos recién egresados- y después de un par de años, los apoyaba para que salieran a hacer un máster. Una vez terminado, ellos tenían que volver a la EI, y desempeñar funciones académicas regulares. Después de un par de años salían nuevamente, pero a hacer un Doctorado, usualmente a una universidad distinta. Tan sólo los procesos de postulación, traslado e

instalación consumían una cantidad enorme de tiempo. Además, el nivel de entrenamiento y especialización en el período entre el máster y el doctorado no era el adecuado como para hacer investigación. Es claro que varios años de lo que en general debería ser la etapa más activa y creativa de una persona para hacer investigación se perdían en estas actividades no académicas. Estos profesores empezaron a regresar al DCC con un Doctorado a fines de los '80 y principios de los '90.

Retomando mi narrativa personal, en 1992 me uní al DCC de la Escuela de Ingeniería de la PUC. Tuve un doble choque cultural. Por un lado, ingresé como académico de la EI de la PUC, la que tenía y tiene prácticas académicas y administrativas muy distintas de las de unidades académicas científicas. Por el otro, me uní ya plenamente a la comunidad de computación académica de Chile. Esta última también tiene prácticas distintas de las de las comunidades científicas tradicionales.

Definitivamente la actividad académica en mi nuevo Departamento no giraba en torno a la investigación de alto nivel. Los pocos trabajos de investigación que se escribían eran enviados a conferencias y medios de publicación más bien regionales, de bajo nivel de exigencia e impacto. Los profesores llegaban con su Doctorado y, en la mayoría de los casos, la investigación se estancaba, desaprovechando el momentum, el estar en el estado del arte, y la energía juvenil.

El DCC, como todos los otros departamentos de la EI estaba sometido a la exigencia tácita de generar dinero para la EI a través de actividades de extensión. Es de esta manera que la EI suplementaba el sueldo de sus profesores. El que el DCC fuera el departamento más joven de la EI, que necesitara solidificarse y establecer una cultura de investigación, no lo libraba de la exigencia. Sin embargo, hay que decir que no todos los profesores veían en ella una maldición. Por un lado, les permitía generar algunos dineros extras para el bolsillo. Por otro, los profesores que tenían más inclinación hacia las actividades de extensión que hacia la investigación, encontraban una situación conveniente. En todo caso, atribuir la poca actividad en investigación al servicio exterior sería una exageración.

Por mi lado, y por decisión personal, orienté mis actividades hacia la investigación, lo que incluía la invitación de investigadores extranjeros y la formación de un grupo de investigación con alumnos de excelencia. También comencé a participar activamente en la Sociedad Chilena de Computación (SCCC). En particular, impulsé la creación de las Jornadas Chilenas de Computación de la SCCC (La Serena, 1993), las que incluyeron el I Encuentro Chileno de Computación, del cual fui el organizador, y la tradicional Conferencia Internacional de la SCCC. Tenía claro que esta última no estaba cumpliendo el rol de acercar a los académicos de la computación en Chile, de permitirles conocerse, intercambiar ideas, y producir relaciones de colaboración y mentoría en investigación. Estos fueron los propósitos detrás de la creación del Encuentro. En 1993 fui elegido miembro del directorio de la SCCC, y tomé el puesto de secretario.

En el terreno de la investigación, comencé, a principios de los '90, a trabajar en la aplicación de formalismos de representación de conocimiento a la especificación de la dinámica de bases de datos relacionales. También a formular y resolver diversos problemas que surgen al razonar sobre la evolución de éstas, por ejemplo, sobre la



1^{as} Jornadas Chilenas de Computación (1993).

satisfacción de restricciones de integridad dinámicas o la evolución de vistas. Esto me llevó a irme metiendo paulatinamente en temas más propios o tradicionales de manejo de datos.

Hacia fines de los 90 comencé a investigar temas de manejo de inconsistencia en bases de datos relacionales. Esto se convirtió en todo un programa de investigación de largo alcance y cobertura. Ahí también caen temas de integración de datos e intercambio de datos entre pares. Esta investigación ha sido siempre de naturaleza científica, con los lenguajes, conceptos y métodos de la matemática, y de la lógica matemática, en particular. Esto no es sorprendente, ya que esta última está en la base misma de las bases de datos relacionales y sus extensiones.

En 1993, antes de que mi período de dos años de prueba en la El terminara, el Decano me pidió que asumiera como Jefe de Departamento. Obviamente, dada mi situación de condicionalidad, ésta era una oferta que no podía rehusar, pero que estaba muy lejos de mis planes y aspiraciones. No hay espacio aquí para describir los tres años de jefatura departamental y los que siguieron inmediatamente a continuación. Eso daría para un artículo completo, o un libro.

Hacia fines del año 2000, descontento, desesperanzado y pesimista sobre mi ambiente de trabajo, y dada la reciente

partida de mi colega Javier Pinto como investigador a Bell Labs en New Jersey, por similares razones y sentimientos, decidí comenzar a buscar alternativas fuera de Chile. Mi deseo era encontrar un ambiente propicio para concentrarme principalmente en investigación.

En Chile siempre hay distracciones inevitables, usualmente en torno a política académica, tanto en el sentido pequeño como en el amplio. En Chile es importante aceptar estas tareas, y creo que contribuí bastante en este terreno. Sin embargo, me parecía que ya era hora de que otros continuaran corriendo la posta.

Dada mi cercanía académica con Canadá, postulé a algunas de sus universidades. En febrero de 2001 recibí la invitación a una entrevista en la Carleton University de Ottawa, la que había visitado doce años antes. Me hicieron una oferta conveniente, y renuncié a la PUC a mediados de 2001. Desde Canadá he intentado seguir contribuyendo al desarrollo de la Ciencia de la Computación en Chile.

Rescato como lo mejor en lo académico de esos años entre 1991 y 2001 mi amistad y colaboración con Javier Pinto. Con él tuve, lejos, la mayor cercanía personal, académica y científica. Su muerte accidental y prematura en el 2001 fue un terrible golpe para mí. Valoro también el contacto con excelentes alumnos que tuve en ese

período, en el cual varios de ellos hicieron o comenzaron a hacer investigación o un posgrado conmigo. Varios de ellos siguieron una carrera académica, y actualmente son investigadores establecidos y reconocidos a nivel internacional. Esto es lo que me da, lejos, la mayor de las satisfacciones. También valoro enormemente las actividades desarrolladas en la SCCC, y las buenas amistades que hice dentro de ella. En dos períodos fui el presidente de la SCCC.

EN UNA COMUNIDAD DE COMPUTACIÓN

Como toda comunidad, la de computación (haciendo la salvedad de que me refiero a la académica) tiene sus peculiaridades, sus prácticas, rituales y personajes. Para mí es inevitable el compararla con otras comunidades científicas y académicas, especialmente con aquellas más tradicionales. Por supuesto, en esta comparación no hay que olvidar que la computación como disciplina científica se inició en los años '30, y la más técnica, aún más tarde. Sin embargo, tampoco es claro que la comunidad de computación vaya a llegar a ser como las más tradicionales, y de que es sólo cosa de darle más tiempo.

Sin pretender hacer un análisis profundo ni detallado, sólo quisiera destacar, en forma muy general, algunas cosas que me llaman particularmente la atención, a modo de reflexión. Ojalá estas observaciones aporten elementos de discusión. Los juicios a continuación se refieren, en general, a la computación como disciplina académica a nivel internacional.

En primer lugar, la cantidad de publicaciones en el área es enorme, excesiva. Y los estándares de calidad son más bien bajos. Más que a la relevancia, es decir el interés o utilidad que una investigación puede tener para el resto de la comunidad, me refiero a los métodos de investigación, a la evaluación de la investigación, y a la novedad de los resultados de la investigación. En el mismo sentido, hay una sobreabundancia de medios de publicación de resultados de investigación. Me atrevo a afirmar que el

que quiera publicar algo va encontrar un medio de publicación, independientemente de los resultados, y casi independientemente de la calidad de la presentación.

Aparte de la sobreabundancia de publicaciones, está el problema del estilo de presentación de la investigación. En los últimos años he participado en muchos comités de programas de conferencias, de reputación decente para arriba, y, en la mayoría de los papers, percibo serios problemas. Usualmente estos no plantean claramente qué es lo que se está proponiendo ni cuáles son los problemas que están siendo atacados. Particularmente ausente está una clara formulación, al principio del paper, de los resultados obtenidos. Abunda la falta de rigor y el tratamiento superficial de los temas. Esto también se manifiesta en una pobre revisión bibliográfica y de resultados de otros autores. Se tiende a rehacer mucha investigación. Sin embargo, hay que reconocer que tanta publicación en la comunidad hace difícil saber qué se está haciendo o se ha hecho.

Aquellos papers de naturaleza o contenido experimental rara vez se preocupan de la replicabilidad de sus experimentos. Muchas veces se llena el paper de gráficos y datos, sin que sea clara su relevancia. Los datos y las condiciones experimentales son rara vez proporcionados.

La comunidad publica mucho en conferencias, mucho más que en journals. Se considera, promueve y defiende esta práctica como algo casi esencial a la investigación en computación. No veo la rapidez del desarrollo y el dinamismo de la computación, usualmente considerados como causa de la preferencia por conferencias, como buenas justificaciones para la enorme cantidad de conferencias y workshops de bajo nivel donde se presentan resultados de investigación. Una cantidad enorme de workshops y conferencias aceptan trabajos mediocres, presentados por investigadores de bajo nivel, ante un público de similar

nivel. Casi no hay discusión ni preguntas. Me cuesta entender la razón y el propósito detrás de la organización de esos encuentros que poco aportan.

Una gran cantidad de trabajos publicados presenta ideas a medio desarrollar. Otros tantos, son papers "visionarios", que muestran cómo van a ser las cosas o como deberían ser, pero sin contribuir a la materialización de esa visión. No podemos dejar de mencionar la enorme redundancia y alto grado de intersección en las publicaciones, entre investigadores y a nivel de investigadores individuales. Hay tanto investigador que empaqueta la misma idea o técnica en las más diversas maneras, incrementando así considerable y artificialmente su número de publicaciones. La comunidad de computación tiene que limpiar sus procesos de publicación.

Para contrapesar esta percepción poco favorable, hay que mencionar que hay también excelentes journals y conferencias, pero son excepcionales, en el sentido estricto del término. Es decir, no es que en computación no haya altos estándares de investigación a los cuales uno pueda hacer referencia. Ha habido excelente investigación y publicaciones en el área desde los comienzos. En consecuencia, hay referentes de calidad. Así es que el problema de la sobreabundancia de publicaciones de bajo nivel no es un problema relacionado con la juventud del área. Es más bien una mala práctica que lamentablemente se ha perpetuado.

Es también importante aclarar que la investigación de alto nivel se publica en conferencias también de muy alto nivel; y que es prestigioso, importante y visible tener papers en estas últimas. Esto sí es característico de la computación y debe ser reconocido. En general, la investigación publicada en estas conferencias no tiene razones para sonrojarse al ser comparada con la mejor investigación realizada y publicada en journals por otras comunidades.

Cada área de computación tiene sus conferencias principales, las de mayor calidad. Estas son usualmente conferencias generales del área, y no conferencias o talleres especializados. Todo investigador en un área sabe (o debería saber) cuáles son las de su área. Un verdadero protagonista en esa área de investigación tiene papers en ellas, es invitado a hacer presentaciones en ellas, y a ser miembro de sus comités de programa. Si no, sólo se está en un papel secundario, o simplemente, como extra.

La comunidad chilena de computación debe apuntar, y también cada investigador local desde el comienzo de su carrera, a publicar en conferencias internacionales del más alto nivel posible. No creo que uno se entrene publicando en conferencias mediocres, locales o regionales. Mas bien, creo que los vicios y malas prácticas se perpetúan y legitiman.

Por esta razón, siempre he pensado que es negativo tener actas formales para conferencias como el Encuentro Chileno de Computación, o la misma Conferencia Internacional de la SCCC. Esta última ya no es tan internacional y, en su formato original ya quedó obsoleta, como todas las conferencias generales de computación. En mi opinión, el Encuentro Chileno de Computación y las Jornadas de la SCCC deben seguir existiendo, pero el énfasis tiene que ser puesto, más que en la publicación de papers y actas, en la oportunidad de intercambio, de mentoría, y de tener buenos tutoriales y conferencias invitadas.

Estas últimas afirmaciones pueden parecer paradójicas a aquellos que me han visto promover el Alberto Mendelzon Workshop on Foundations of Data Management. Éste se está realizando anualmente en América Latina. Mi visión del Workshop es que éste sea de alta calidad y prestigio internacional, con los más altos estándares, sólo que realizado en nuestra región, para beneficio de sus investigadores y estudiantes.

Un segundo aspecto de la comunidad que me ha llamado la atención es que ésta es muy generosa y tolerante con sus miembros en lo que se refiere a la calidad de la investigación y a la ausencia de ella. Una parte importante de la producción en investigación es de nivel más bien bajo. Hay un alto número de investigadores en cada área de la computación que hacen una carrera basada en publicaciones en conferencias de esa área que son de dudosa calidad en términos de nivel de exigencia e impacto.

Otras comunidades son mucho más estrictas con lo que se considera investigación y publicación; y con el nivel que debe alcanzar en ella un académico para ser promovido a una categoría superior o, incluso, para mantener su lugar en la universidad (cuando no está la seguridad de un "tenure"²). En este respecto, excepto en un caso, los departamentos de computación en Chile no tienen niveles de exigencia muy altos.

Existe en cierto modo el mito, en todas partes, de que no se hace tanta investigación porque los académicos hacen computación, léase ingeniería, y aplicaciones y desarrollos tecnológicos. Yo no he percibido esto último como la causa. Son pocos los que realizan estas actividades. Y en la mayoría de esos pocos casos, el nivel de innovación tecnológica es más bien bajo. No toda la actividad técnica que hace

un académico, por el sólo hecho de ser realizada por un académico, es investigación. Muy probablemente, tampoco desarrollo tecnológico innovativo ni ingeniería del más alto nivel.

En la comunidad abundan los investigadores que desarrollan una carrera sobre la base de las aplicaciones de la computación. Llamemos C a la vasta área que incluye a la computación en sus diversas modalidades (Ciencia de la Computación, Ingeniería de Computación, Tecnología de la Información, Ingeniería de Software, etc.); y E, a un área distinta, genérica, donde se puede aplicar C. Hay muchos papers que presentan como resultados de investigación ciertas aplicaciones de C en E.

Hay varias preguntas que surgen naturalmente en ese escenario: (a) ¿Es eso investigación o la aplicación de una tecnología ya establecida en C? Si es lo segundo, entonces no es investigación desde el punto de vista de C, sino más bien ingeniería. (b) Tal vez aceptando que es investigación, ¿es eso investigación en C o en E? (c) Mirando más en detalle, ¿es esa investigación relevante para C o para E (o ninguna de las anteriores)?

Debo reconocer que veo a ese tipo de investigación con cierto grado de suspicacia. Si hay alguna contribución a C en esas aplicaciones, el investigador radicado en C

debería ser capaz de identificar y destilar lo que es relevante para C y publicarlo como un avance para C. Por otro lado, si desde el punto de vista de C es sólo la aplicación de una tecnología establecida, entonces son los miembros de E los que deberían juzgar y valorar esa investigación. Por supuesto, el juicio de estos últimos podría dejar legítimamente indiferentes a los de C.

Puesto de otra manera, a no ser que una institución haya decidido deliberadamente desarrollar el área de "aplicaciones de C en E", o que sólo se interese en contar papers, el departamento de C tiene el derecho a analizar y valorar con especial precaución este tipo de investigación. Este síndrome de las aplicaciones interdisciplinarias se da naturalmente en computación, pero otras disciplinas también lo comparten, por ejemplo, estadística. Aplicaciones de computación son importantes, pero deben ser valoradas como tales, y no necesariamente como investigación o publicación.

Finalmente, como tercer punto, está un tema que siempre me ha preocupado, y sobre el cual trato de alertar a los futuros estudiantes de doctorado. La formación que puede recibir un joven a través de sus estudios de posgrado varía mucho dependiendo del estilo de trabajo de su supervisor o institución. En computación abunda el profesor que arma un enorme equipo de investigación, del cual se convierte más en un gerente que en un miembro activo en la investigación directa. El alumno pasa a ser una pieza dentro de ese megaproyecto, muchas veces de desarrollo de algún software o aplicación. Con frecuencia, el rol del alumno se convierte en el de implementar una parte. Si, además, "diseña" lo que implementó, ese diseño se considera parte de su investigación. Así puede obtener finalmente su grado.

Es probable que esa actividad haya significado enorme trabajo, pero no es claro que ese alumno haya adquirido una buena



Inauguración 1^{as} Jornadas Chilenas de Computación (1993).

² No puedo suponer que todos los lectores saben lo que es el *tenure*. Este es el status otorgado por una universidad en Estados Unidos o Canadá a un profesor. Éste implica que ese profesor no puede ser despojado de su puesto por razones de pobre productividad académica. Sólo se puede perder por razones más o menos obvias, no estrictamente académicas, o por reestructuración de la unidad académica. Típicamente uno es contratado como profesor asistente sin *tenure*. Después de unos seis años es evaluado. Ahí recibe el *tenure* y es promovido a profesor asociado o tiene que dejar la universidad.

formación. No es claro que el alumno haya tenido una visión nítida y global de lo que se estaba haciendo ni que haya quedado bien, críticamente y operacionalmente informado del estado del arte en su área y de la literatura relevante. Es muy probable que no haya desarrollado las cualidades de autonomía, independencia y autocrítica que si se alcanzan a través del desarrollo de una tesis en una tema que se pare por sí mismo, que sea realizable por el alumno por entero, de principio a fin, incluyendo la formulación precisa y circunscrita del proyecto a desarrollar. He conocido a muchos doctores que nunca aprendieron a investigar por este motivo. Ésta es una nota de precaución para los jóvenes que aspiran a hacer un doctorado. Por supuesto, este problema no se da sólo en computación, sino en cualquiera otro donde se puedan generar grandes implementaciones o aplicaciones.

Es imposible no relacionar estos comentarios sobre el Doctorado con la etapa posiblemente más difícil en la carrera de un investigador. Precisamente aquella que viene inmediatamente después de terminar el Doctorado. Es ahí cuando hay que empezar a comportarse, al menos en materias de investigación, como un ser inteligente, independiente, conocedor y creativo. El tipo de Doctorado que uno haya hecho tiene fuerte influencia en el éxito en este período. Sin embargo, aunque no se haya hecho uno de la mayor calidad o en las óptimas condiciones, por diversas razones, desde personales hasta académicas, siempre es posible compensar y recuperarse inmediatamente después. En ese momento es importante buscar alianzas en investigación o franca mentoría de parte de investigadores sólidos y establecidos, dentro y fuera del país. Es clave en esa etapa nunca perder el momentum ni parar la actividad de investigación. Hay que trabajar duro y concentradamente, sin distraerse con otras actividades, especialmente administrativas y de extensión. Hay que publicar, apuntando alto, y postular a proyectos de investigación. Un vacío en investigación en esa etapa es peligroso y difícil de remontar.

Todo esto es particularmente importante en Chile, donde no existe, en general, el sistema de *tenure*. Los recién doctorados son contratados en la universidad y no corren peligro de perder su trabajo si no son productivos en investigación. Mi consejo es establecerse en Chile recién Doctorado adoptando la mentalidad del *"tenure-track"*, es decir, como si uno fuera a ser evaluado en investigación después de cinco o seis años, cuando se decidiría si se le otorga el *tenure*.

Las instituciones chilenas que acogen a los recién doctorados deben proveer las condiciones para que esa etapa inicial sea productiva, y, a la vez, deben exigir que haya un nivel alto de productividad. Esta última debe ser medida en función del número y la calidad de las publicaciones, y en el grado de éxito en concursos competitivos de proyectos de investigación, especialmente aquellos que reflejan en cierto modo, el nivel alcanzado en investigación, como los Fondecyt.

El doctorado es una experiencia en investigación. Es una forma de aprender a hacer investigación haciéndola bajo supervisión. No hay justificación para que todo ese esfuerzo, individual y a veces institucional, se pierda. Una institución que acoge a un profesor recién doctorado, a no ser que sólo quiera hacer propaganda exhibiendo una serie de nombres con doctorado, debería hacer todo lo posible para que éste dé frutos en investigación, como se esperaría. Después de todos estos años, todavía creo que lo que mejor se aprende es aquello que se aprende a través de la investigación. También creo que lo que mejor se enseña, al menos a nivel académico, es aquello que es iluminado desde la perspectiva de la investigación.

EN CONCLUSIÓN

Mi motivación detrás de las críticas que hice más arriba es la de alertar sobre ciertas prácticas que considero negativas. Es claro que no vamos a cambiar la comunidad mundial de computación, pero sí podemos

introducir algunos cambios en nuestro entorno, e impedir o minimizar el caer en los errores de otros.

Mi propia investigación es, sin duda, en Ciencia de la Computación. Como investigador en computación, mis críticas las entrego desde el interior de la misma comunidad a la cual pertenezco, porque me importa y me siento comprometido con ella. Las comparto con mi comunidad chilena porque, en Chile, la computación académica es aún joven, y se puede beneficiar con una discusión con altura de miras en torno a estos temas. Este tipo de discusiones debería ayudar a fortalecer y a consolidar una comunidad científica, a subir los niveles de calidad de la investigación científica, y a hacer surgir figuras en ella, muy necesarias, que tengan el más alto nivel de excelencia en investigación y un considerable peso académico e intelectual, tanto dentro como fuera de la comunidad.

Algunos de los problemas que indique en la sección anterior son muy probablemente compartidos con otras comunidades de investigación, no son propiedad de computación. La mezcla de ciencia y tecnología que se encuentra en computación también aparece en otras áreas. Esta mezcla complica la evaluación de la investigación, pero a la vez enriquece a esta última.

Las críticas son relevantes cuando hay potencial y esperanzas. Dentro de América Latina, Chile presenta un excelente nivel en investigación en computación. Estamos primeros o segundos, dependiendo de cómo se midan las cosas. Da orgullo y satisfacción el ver que un sólido y creciente número de investigadores chilenos está jugando, desde Chile, en las primeras ligas internacionales. Esto se ve más claramente en las áreas de investigación en torno a teoría y algoritmos, a manejo de datos, y a la Web, pero hay otras áreas donde también se está empezando a generar investigación relevante e internacionalmente apreciada. Este éxito debe ser reconocido y apoyado, para multiplicarlo, y para el beneficio del país. BITS



Programas de posgrado chilenos en Ciencia de la Computación e Informática: un diagnóstico inicial



Sergio F. Ochoa

Ingeniero de Sistemas, Universidad Nacional del Centro de la Provincia de Buenos Aires (UNICEN), Argentina; Doctor en Ciencias de la Ingeniería, mención Computación, Pontificia Universidad Católica de Chile. Profesor asistente del Departamento de Ciencias de la Computación, Universidad de Chile. Sus áreas de interés son: ingeniería de software y sistemas colaborativos. Además es miembro del Joint Steering Committee de LACCIR y representante chileno ante CLEI.
sochoa@dcc.uchile.cl



Marcello Visconti

Ingeniero Civil Informático de la Universidad Técnica Federico Santa María, Valparaíso, Chile, y Ph.D. in Computer Science de Oregon State University, EE.UU. Académico del Departamento de Informática de la Universidad Técnica Federico Santa María, donde desarrolla investigación y docencia en ingeniería de software, calidad de software y mejoramiento de procesos de software.
visconti@inf.utfsm.cl

Hay diversos motivos por los cuales la formación de posgrado, especialmente en áreas tecnológicas, se ha vuelto una necesidad cada vez más tangible. Una razón importante radica en la velocidad a la que avanza el conocimiento en estas áreas, que hace que profesionales y académicos requieran actualizar periódicamente sus conocimientos. Por otra parte, durante los últimos años ha habido en la industria un fuerte énfasis hacia la innovación, como palanca de los esfuerzos nacionales por alcanzar el pleno desarrollo del país en los próximos años. Esto implica, en parte, poder contar con capital humano capaz de hacer investigación y transferir dichos resultados al aparato productivo, de modo de materializar la necesaria innovación que asegure el logro de los objetivos nacionales de desarrollo. Estas capacidades de investigación e innovación han sido históricamente parte de la formación que los alumnos reciben en los programas de posgrado. Por lo tanto, cada día más y más profesionales de la industria se embarcan en este tipo de programas.

Finalmente, también está la demanda histórica de programas de posgrado, la cual está vinculada a la academia. Muchas universidades buscan mejorar la formación de sus académicos, y volverse así más competitivas tanto en el ámbito docente como de investigación, innovación y transferencia tecnológica. Para ello envían a sus docentes a hacer posgrados, generalmente fuera de la institución. La principal razón de esto radica en el hecho de que un posgrado es bastante más que el conocimiento científico/técnico que el alumno adquiere. Las vivencias y el conocimiento acerca de los procesos que se llevan a cabo en otras universidades, retroalimentan al sistema una vez que el alumno regresa a su universidad de origen.

Varios son los programas de posgrado chilenos en el área de Ciencias de la Computación e Informática, que se han ganado el reconocimiento de estudiantes y profesores, tanto dentro como fuera del país. Por esa razón, año tras año alumnos de diversos países (fundamentalmente de

Sudamérica) llegan en busca de nuevo conocimiento y oportunidades para realizar investigación. La cantidad de personas aceptadas en programas de posgrado es usualmente bastante superior a la que efectivamente ingresa a los mismos. A pesar de que los alumnos han pasado un filtro exigente para ser aceptados en un programa, la principal limitante son los recursos económicos requeridos para llevar a cabo sus estudios. El sistema de becas de posgrado chileno está enfocado principalmente en apoyar a alumnos chilenos que realizan estudios de doctorado. Sin embargo, las necesidades asociadas a los procesos de innovación en la industria requieren profesionales mayoritariamente con estudios de magíster.

Tres tipos de programas actualmente responden a la demanda de estudios de posgrado en Computación e Informática en Chile: Doctorados, Magíster en Ciencias y Magíster Profesionales. Cada uno de ellos tiene un nicho bien definido y cumple una función específica. A continuación se describe cada una de estas categorías y se presenta la nómina de los programas más conocidos.

PROGRAMAS DE DOCTORADO

Los programas de Doctorado chilenos son particularmente atractivos para alumnos de la región, especialmente para gente que trabaja en ambientes académicos, o que pretende hacerlo en un futuro cercano. Estos programas están enfocados en la investigación y por lo tanto tienen un fuerte componente teórico. Su duración es de cuatro años con dedicación full time, pero en la práctica es un poco más que eso. En Chile hay sólo cinco programas de Doctorado, los cuales cuentan con un importante cuerpo académico que los respalda. La Tabla 1 muestra la nómina de programas ordenados por fecha de creación.

En la Tabla 1 se indica con (*) aquellos programas que están acreditados por la Comisión Nacional de Acreditación de Chile (CNA). Esto implica que los estudiantes de dichos programas pueden optar a una beca del sistema de becas chileno.

Tabla 1
Programas de Doctorado en Computación e Informática

Año de Creación	Universidad	Nombre del Programa
1990 (*)	Pontificia Universidad Católica de Chile (PUC)	Doctorado en Ciencias de la Ingeniería, área de especialización Ciencia de la Computación
1997 (*)	Universidad de Chile (UCH)	Doctorado en Ciencias mención Computación
2000	Universidad de Santiago de Chile (USACH)	Doctorado en Ciencias de la Ingeniería mención Informática
2003 (*)	Universidad Técnica Federico Santa María (UTFSM)	Doctorado en Ingeniería Informática
2010	Universidad de Concepción (UdeC)	Doctorado en Ciencias de la Computación

En promedio, a estos programas ingresan anualmente 23,5 alumnos y egresan alrededor de diez, lo cual implica que la tasa de egreso promedio es de 43% aproximadamente. Sin embargo cabe hacer notar que la dispersión en la tasa de graduación es muy alta entre los distintos programas. De los alumnos ingresantes, sólo el 20 % de ellos son extranjeros. Sin embargo, la dispersión en este ámbito es sumamente importante entre las distintas universidades. La Tabla 2 muestra la información de ingresantes, egresados, tasa de graduación y porcentaje de extranjeros, desglosada por programa.

PROGRAMAS DE MAGÍSTER EN CIENCIAS

Los programas de Magíster en Ciencias guardan alguna similitud con los Doctorados; sin embargo son mucho más cortos y menos

profundos. Si bien también están orientados a la investigación, su duración es de dos años aproximadamente.

Este tipo de programa fue inicialmente pensado para académicos que querían llevar a cabo sus labores docentes y de investigación en una universidad, sin tener que pasar por un Doctorado. Sin embargo, los parámetros por los que se mide hoy a las Instituciones de Educación Superior, hacen que las mismas se vuelquen más hacia los Doctorados que hacia los Magíster en Ciencias. Los Magíster en Ciencias siguen cumpliendo en parte ese rol, aunque se los ve más bien como un paso intermedio hacia un futuro Doctorado.

Son relativamente pocos los programas de Magíster en Computación e Informática en el país. La Tabla 3 muestra la nómina de los mismos por fecha de creación; sólo cuatro de ellos están acreditados por la CNA (marcados con *).

Tabla 2
Promedio Anual de Ingresantes/Egresados de Programas de Doctorado

Indicador / Programa	Doct. en Cs. de la Ing. m/ CC. - PUC (1990*)	Doct. en Cs. m/ Computación - UCH (1997*)	Doct. en Cs. de la Ing. m/ Informática - USACH (2000)	Doct. en Ing. Informática - UTFSM (2003*)	Doct. en Cs de la Computación - UdeC (2010)	TOTAL
# Alumnos Ingresantes	10	6	2	3,5	2	23,5
# Alumnos Egresados	7	2,5	0,5	0,1	N/A	10,1
Tasa de Graduación	70%	42%	25%	3%	N/A	43%
% Alumnos Extranjeros	50%	43%	< 5%	< 5%	0%	~20% (+)

(+): Representa el valor promedio entre los cinco programas.

En general se ven muy pocos profesionales de la industria en este tipo de programas, excepto en ciudades donde la única opción de posgrado en el área es un Magíster en Ciencias. En general los profesionales buscan programas más aplicados a su quehacer diario; rol que normalmente juegan los Magíster de corte profesional.

Anualmente ingresan a los programas de Magíster en Ciencias un promedio de 93 personas y egresan 58, lo cual da una tasa de graduación del 61%. El número promedio de extranjeros que ingresa a estos programas es de alrededor del 9% del total de alumnos ingresantes. La tabla 4 muestra la información de ingresantes, egresados, tasa de graduación y porcentaje de extranjeros, desglosada por programa. Los programas que se presentan en dicha tabla son aquellos de los cuales se pudieron obtener los datos para este estudio.

PROGRAMAS DE MAGÍSTER PROFESIONAL

Este tipo de programa es relativamente nuevo en el área de Computación e Informática en el país. Sin embargo su crecimiento en términos de cantidad de alumnos ha sido un tanto explosivo. Estos programas buscan llevar la teoría a aspectos prácticos del quehacer profesional del estudiante. Su duración es de aproximadamente dos años, se dictan en horarios vespertinos y requieren dedicación part time, lo que permite a los profesionales participar del programa luego de su jornada laboral.

Los programas consideran un trabajo de tesis o tesina, la cual generalmente está orientada a resolver un problema de la industria. Allí los alumnos deben aplicar los conocimientos adquiridos durante el programa, a fin de construir una solución innovadora a un problema relevante en su ámbito de desarrollo profesional. En general este tipo de tesis o tesina requiere también un poco de investigación, aunque esencialmente se trata de aplicar el conocimiento ya adquirido. Un aspecto muy interesante de esta tesis o tesina, radica en el hecho de que se recomienda al alumno realizar dicho trabajo final en su ámbito laboral, bajo la supervisión de un académico del programa. De esa manera, tanto el alumno

Tabla 3
Programas de Magíster en Ciencias

Año de Creación	Universidad	Nombre del Programa
1975 (*)	Universidad de Chile	Magíster en Ciencias mención Computación
1982 (*)	Pontificia Universidad Católica de Chile	Magíster en Ciencias de la Ingeniería mención Computación
1990 (*)	Universidad Técnica Federico Santa María	Magíster en Ciencias de la Ingeniería Informática
1994	Universidad de Santiago de Chile	Magíster en Ingeniería Informática
1994 (*)	Universidad de Concepción	Magíster en Ciencias de la Computación
2001	Pontificia Universidad Católica de Chile	Magíster en Ingeniería
2002	Universidad de Tarapacá	Magíster en Ingeniería de Software
2006	Universidad Católica del Norte	Magíster en Ingeniería Informática
2006	Pontificia Universidad Católica de Valparaíso	Magíster en Ingeniería Informática
2009	Universidad del Bío-Bío	Magíster en Ciencias de la Computación
2009	Universidad Diego Portales	Magíster en Ciencias de la Ingeniería mención Informática y Telecomunicaciones

Tabla 4
Promedio Anual de Ingresantes/Egresados de Programas de Magíster en Ciencias

Indicador/Programa	Mag. en Cs. m/Comp. - UCH (1975*)	Mag. en Cs. de la Ing. m/CC. - PUC (1982*)	Mag. en Cs. de la Ing. Inf. - UTFSM (1990*)	Mag. en Ing. Inf. - USACH (1994)	Mag. en Cs. de la Comp. - UdeC (1994*)	Mag. en Ing. Inf. - PUCV (2006)	Mag. en Cs de la Comp. - UBío-Bío (2009)	TOTAL
# Alumnos Ingresantes	14	15	10	17	10	18	9,5	93,5
# Alumnos Egresados	7	12	6	11,6	7,2	13	N/A	56,8
Tasa de Graduación	50%	80%	60%	68%	72%	72%	N/A	61%
% Alumnos Extranjeros	15%	20%	< 5%	< 5%	10%	2%	10%	~9% (+)

(+): Representa el valor promedio.

como la institución a la que él pertenece se ven beneficiados con su realización.

Pese a la poca experiencia que aún se tiene con este tipo de programa, hasta el momento han demostrado ser una interesante alternativa para generar sinergia entre la academia y la industria. Parte de la innovación que requiere llevar a cabo la industria, podría venir de trabajos de tesis o tesina de alumnos de este tipo de programas. Debido a la orientación de los Magíster Profesionales, el cuerpo académico de los mismos usualmente mezcla gente

de la industria y de la academia. La Tabla 5 muestra la nómina de programas de este tipo, en el ámbito de Computación e Informática en Chile. Por el momento hay sólo un Magíster Profesional acreditado, sin embargo hay varios otros que están haciendo esfuerzos para lograr su acreditación en el corto plazo.

Es de esperar que en los próximos años se amplíe la oferta y la demanda por este tipo de programas, tanto en Santiago como en regiones. A partir del trabajo realizado por las primeras generaciones de graduados de

estos programas, las organizaciones han comenzado a ver una ventaja concreta y tangible en el hecho de financiar parte de los estudios de Magíster Profesional de su personal. Sin embargo para que esta ventaja se mantenga en el tiempo, es importante que este tipo de programas mantenga un nivel académico adecuado. Sólo de esa manera se puede esperar que crezcan en forma saludable y sostenible.

En lo que respecta a la demanda por Magíster Profesionales, la Tabla 6 muestra el detalle de ingresantes, egresados, tasa de graduación y porcentaje de extranjeros por programa. Los programas que se presentan en dicha tabla son aquellos de los cuales se pudieron obtener los datos para este estudio.

Comparando los números de las Tablas 4 y 6, que corresponden a los Magíster en Ciencias y Magíster Profesionales respectivamente, se puede ver que los totales y los promedios son similares. Sin embargo, es importante hacer notar que estamos comparando siete programas de Magíster en Ciencias contra tres Magíster Profesionales. Los resultados muestran a las claras la preferencia de la gente por los programas de corte profesional.

CUERPOS ACADÉMICOS

El cuerpo académico que está detrás de los programas juega un rol fundamental en la calidad y en la evolución de los mismos. La Tabla 7 muestra un resumen de los investigadores de jornada completa que están vinculados a las instituciones que imparten los programas de posgrado antes mencionados.

En la tabla anterior se puede ver que en casi todas las instituciones, el número de investigadores con Doctorado está por encima de la cantidad de personas con Magíster. Esto no es demasiado sorprendente ya que uno de los indicadores por los cuales se mide a un programa (inclusive los de pregrado), es el máximo grado obtenido por los miembros de su cuerpo académico. En la tabla se puede apreciar también que el número de investigadores de jornada completa asociados a cada una de estas instituciones es distinto, sin

Tabla 5
Programas de Magíster Profesional

Año de Creación	Universidad	Nombre del Programa
2004 (*)	Universidad de Chile	Magíster en Tecnologías de la Información
2004	Universidad Técnica Federico Santa María	Magíster en Tecnologías de la Información
2007	Pontificia Universidad Católica de Chile	Magíster en Tecnologías de la Información y Gestión
2009	Universidad Andrés Bello	Magíster en Informática
2009	Universidad Católica del Norte	Magíster en Gestión de Información y Tecnologías
2011	Pontificia Universidad Católica de Chile	Magíster en Gestión de la Información y Bibliotecología

embargo en la mayoría de los casos este número cae dentro del rango de "10 a 25 investigadores". El porcentaje de mujeres entre estos investigadores es en general bajo, sin embargo hay casos excepcionales donde este número es alrededor del 30%.

DIAGNÓSTICO INICIAL

La mayoría de los programas de posgrado en Computación e Informática son competitivos, y se vuelven más y más competitivos con el paso del tiempo. El proceso de acreditación que lleva a cabo la CNA ayuda en forma sistemática a lograr este objetivo. Otro elemento que contribuye a su competitividad, es la reciente incorporación de muy buenos académicos jóvenes, tanto chilenos como extranjeros, por parte de diversas universidades del país. Estos jóvenes que se incorporan al sistema traen un nuevo empuje e ideas para implantar en nuestros programas de posgrado. La Figura 1 muestra la evolución del número de programas de posgrado, por categoría, desde el inicio en 1975.

Desde 1975 al 2000 hubo poca actividad en términos de creación de programas de posgrado en el área. Sin embargo desde el 2000 a la fecha la oferta ha crecido en forma notable, tanto en la capital como en regiones. Claramente la pendiente de crecimiento del número de programas de Magíster de corte profesional, muestra a las claras que este tipo de posgrado será protagonista durante los próximos años. Una razón importante de dicho crecimiento radica en el hecho de que el ingreso de alumnos a estos programas es menos sensible a la disponibilidad de becas. Otra razón de crecimiento parece ser su orientación, puesto que estos acogen de mejor manera los intereses por perfeccionamiento de una proporción mayoritaria de los profesionales, quienes son los potenciales interesados en cursar un Magíster Profesional.

El crecimiento de los otros programas, especialmente los Doctorados, seguirá dependiendo en gran medida del número de becas disponible para sus estudiantes, puesto que demandan dedicación full time. Hoy las "becas de doctorado para

Tabla 6.
Promedio Anual de Ingresantes/Egresados de Programas de Magíster Profesionales

Indicador / Programa	Magíster en TI - UCH (2004*)	Magíster en TI - UTFSM (2004)	Magíster en TI y Gestión - PUC (2007)	TOTAL
# Alumnos Ingresantes	6	50	28	84
# Alumnos Egresados	2	29	19	50
Tasa de Graduación	33%	58%	68%	60%
% Alumnos Extranjeros	40%	< 5%	12%	~19% (+)

(+): Representa el valor promedio.

Tabla 7
Investigadores de Jornada Completa por Institución

Institución	Departamento/Instituto	Total Investig.	Inv. con Doct.	Inv. con Mag.	% de Mujeres
Pontificia Universidad Católica de Chile	Departamento de Ciencia de la Computación	14	14	0	14%
Pontificia Universidad Católica de Valparaíso	Escuela de Ingeniería Informática	13	8	5	8%
Universidad Andrés Bello	Departamento de Informática	5	2	3	0%
Universidad Católica del Norte	Departamento de Ingeniería de Sistemas y Computación	16	7	9	13%
Universidad de Chile	Departamento de Ciencias de la Computación	21	20	1	15%
Universidad de Concepción	Departamento de Ingeniería Informática y Ciencias de la Computación	14	9	5	36%
Universidad de Santiago de Chile	Departamento de Ingeniería Informática	16	13	3	19%
Universidad del Bío-Bío	Departamento de Informática	23	12	11	26%
Universidad Diego Portales	Escuela de Ingeniería Informática y Telecomunicaciones	8	5	3	13%
Universidad de Tarapacá	Área de Computación e Informática	10	7	3	0%
Universidad Técnica Federico Santa María	Departamento de Informática	30	22	8	13%
Total:		170	119	51	14% (+)

(+): Representa el valor promedio.

chilenos” representan la categoría con más apoyo financiero. Sin embargo, el número de personas que logra obtener financiamiento en ese track, está muy por debajo de la demanda.

Si bien el número de personas que ingresa a los programas de posgrado es bajo, bastante más preocupante es el número de estudiantes que egresa de los mismos. En el caso de los programas de Magíster la tasa de egreso es del 60% aproximadamente, y en los Doctorados es de un 42%. Claramente ese es un tema que necesita ser abordado con más fuerza por parte de las instituciones que los imparten. Aparentemente el retraso en la salida de los estudiantes se produce durante la realización de la tesis, sin embargo no queda claro cuál es la causa por la que se genera esta situación.

En términos de investigadores, Chile comienza a manejar una fuerza de investigación interesante, especialmente porque una parte importante de ellos tienen doctorados recientes. La cantidad de investigadores asociados a programas de posgrado está muy por encima de la demanda generada por los alumnos que ingresan a los mismos. Los valores antes presentados muestran que hay 1,2 estudiantes de posgrado por cada

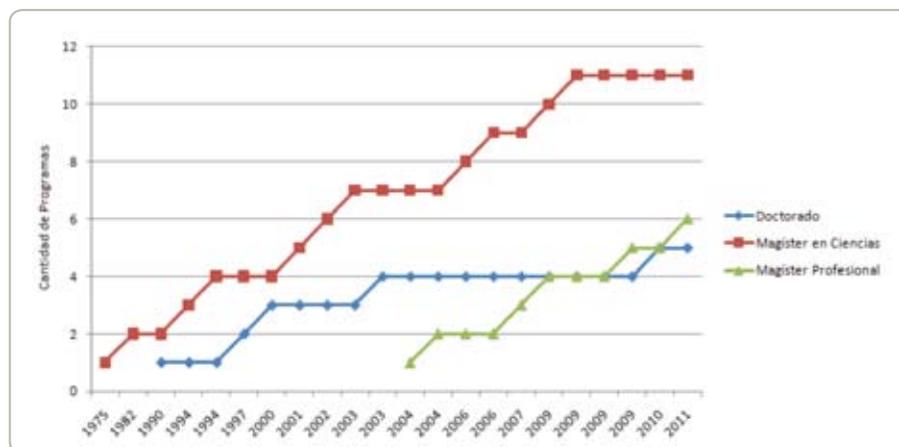
investigador, lo cual indica que el número de estudiantes está muy por debajo de lo razonable.

Independiente de las limitaciones aquí identificadas, los programas de posgrado en Chile parecen gozar de buena salud. El reconocimiento que muchos de ellos tienen a nivel nacional e internacional es una muestra de que no sólo se están haciendo las cosas bien, sino que además se trata de mejorar en forma sostenida. La principal limitante que hoy enfrentan estos programas es la falta de becas que ayuden a que más y más personas puedan acceder a los mismos.

AGRADECIMIENTOS

Los autores desean agradecer a las diversas personas que contribuyeron en la recolección de la información presentada en este artículo; particularmente a los académicos: Gonzalo Acuña y Mauricio Marín (Universidad de Santiago), Yadrán Eterovic y Felipe Guerrero (Pontificia Universidad Católica de Chile), Marcela Varas y John Atkinson (Universidad de Concepción), Silvana Roncagliolo (Pontificia Universidad Católica de Valparaíso), Angélica Caro y Gilberto Gutiérrez (Universidad del Bío-Bío). BITS

Figura 1
Evolución de los Programas de Posgrado en Computación e Informática en Chile



Hardware / Redes / Software
 Lenguajes / Algoritmos /
 Criptografía / Estructura
 de Datos / Sistemas de
 Información / Interacción
 Humano-Computador / Sistemas
 Colaborativos / Sistemas
 Inteligentes / Computación
 Gráfica / Computación Científica

Panorama de la Ciencia de la Computación en Chile

Desde la génesis del presente número de la Revista Bits, nuestra idea siempre fue la de intentar comprender cuál es el estado de la Ciencia de la Computación (CC) en Chile hoy. Sin embargo, existían al menos dos maneras de hacer esto. La primera -más tradicional- consistía en encargarle a un grupo de especialistas analizar el estado de nuestra disciplina en el país en detalle, en particular, los lugares donde se realiza investigación, los académicos que trabajan haciendo investigación en el área, las redes de colaboración existentes entre estos investigadores, entre varias otras. La ventaja que tiene este tipo de mirada es la uniformidad en el tratamiento de los temas, además de la posibilidad de cubrir más “radiográficamente” el espectro de la CC en Chile; es decir, éste podría haber sido un análisis, que de hacerse con el debido cuidado, no hubiera dejado investigador ni investigación sin cubrir.

Sin embargo -a costo de perder en uniformidad y extensión- nos decidimos por una alternativa más dinámica: invitar

a los mismos investigadores (o a grupos de investigación, en caso de que los investigadores participantes se pusieran de acuerdo) a que nos contaran qué hacen, con quién trabajan, dónde publican, etc. Creemos que esto da más frescura a la Revista, pues pensamos que nadie mejor que el investigador mismo para describir con real pasión el trabajo que está realizando. Por supuesto, por el otro lado esta decisión tiene desventajas, entre las cuales las más obvias son la subjetividad en el criterio de elección de los invitados y la posibilidad de que algunos investigadores no respondieran al llamado.

Con respecto a la primera desventaja, la subjetividad en la elección de los invitados, me gustaría decir algo con respecto a cómo se hizo la selección de los investigadores participantes. En primer lugar, quisimos incluir sólo a investigadores activos, es decir, gente que esté participando en investigación en algún área de CC en los últimos años. Aunque no es absolutamente medible, esto es fácilmente ponderable observando



las publicaciones de los investigadores en sitios como DBLP o Google Scholar. Y es importante destacar que todos los invitados a escribir para la Revista pasan ese filtro con creces. Es decir, todos los que están aquí merecen estar aquí. El problema es el opuesto: es probable que no todos los que merecen estar aquí finalmente estén aquí, simplemente porque como editor cometí el error de no invitarlos. Me hago absolutamente responsable de ese error, y pido de antemano las disculpas del caso. Aunque no sirve como atenuante, debo al menos decir que ese error ha sido completamente involuntario, y se explica solamente por la premura con que se recolecta la información de esta Revista. En otras palabras: hemos tratado de ser lo más inclusivos posibles, pero no necesariamente hemos podido ser exhaustivos.

Sobre el segundo problema, la deserción de algunos investigadores a participar en la Revista, es poco lo que podemos decir, salvo que en la mayoría de los casos se debió a exceso de carga académica del invitado. A nosotros nos hubiera encantado mostrar el trabajo de todos los seleccionados, pero lamentablemente no siempre es posible.

Para estructurar los resúmenes de investigación recibidos, dividimos a los investigadores y/o grupos de investigación según áreas. En este caso optamos por la siguiente división: (A) Hardware y Redes, (B) Lenguajes y Software, (C) Algoritmos, Criptografía y Estructuras de Datos, (D) Sistemas de Información, (E) Sistemas Colaborativos e Interacción Humano-Computador, y (F) Sistemas Inteligentes. Esta división intenta seguir del modo más fiel posible las categorías propuestas por la ACM, respetando a la vez la estructura particular de los grupos en nuestro país.

Acerca de qué es posible observar en los artículos recibidos, preferiría dejar de lado un análisis con mucho detalle, y más bien darle la posibilidad a cada lector de sacar sus conclusiones. Sin embargo, sólo a vuelo de pájaro, es posible mencionar que claramente en términos del contexto latinoamericano se observa una comunidad en CC bastante sana y productiva, y en la que es fácil observar que algunos de sus integrantes se hallan en el primer nivel mundial de sus respectivas áreas. En particular, áreas como: Lenguajes, Software, Algoritmos y Estructuras de Datos, Sistemas

de Información y Sistemas Inteligentes, están bastante desarrolladas en nuestro país y cuentan ya con una importante masa crítica de investigadores.

Por supuesto, la muestra que hacemos aquí es, además de parcial, bastante superficial. De hecho, sería interesante no sólo tener una vista cualitativa de nuestra investigación, sino también una vista cuantitativa que describa cosas como cuánto estamos publicando, cuál es el impacto de nuestra investigación, en cuántos comités de programa participamos, y una larga lista de etcéteras. Hacer este tipo de análisis escapa definitivamente a las competencias de nuestra Revista.

Finalizando, me gustaría plantear un tema para el futuro, que apareció al reunir los artículos: ¿Cuál es el nivel de la CC en Chile comparado con respecto a otros lugares en Latinoamérica, y algunos países europeos como España y Portugal? ¿Qué tan lejos estamos de ellos? Y quizá más importante: ¿Cómo nos proyectamos como disciplina y a dónde nos gustaría llegar?

A continuación les presentamos grupos de investigación e investigadores de nuestro país, divididos en las áreas anteriormente descritas.

Profesor Pablo Barceló
Editor Revista Bits de Ciencia

Hardware + Redes

Reinaldo Vallejos:

EVALUACIÓN DE RENDIMIENTO DE SISTEMAS COMPUTACIONALES Y DE REDES DE COMPUTADORES

Departamento de Ingeniería Civil Telemática, Universidad Técnica Federico Santa María.

Mi área de trabajo tiene que ver con evaluación de sistemas computacionales y redes de computadores por medio de modelos Markovianos. Esta área abarca un amplio espectro de problemas de interés, los cuales tienen en común la evaluación cuantitativa de rendimiento de los sistemas. Dentro de los problemas de redes, últimamente el grupo se ha enfocado en el análisis de desempeño de redes ópticas y redes inalámbricas, debido a que las primeras constituyen el backbone de las redes actuales y las segundas son las más usadas por las personas para acceder a los servicios de redes de computadores.

Un tema siempre presente en la investigación es el desarrollo de nuevos métodos matemáticos para evaluar diferentes medidas de rendimiento asociadas a los problemas estudiados, y algoritmos computacionales para implementar éstos métodos. En particular nos interesa resolver cadenas de Markov con un gran número de estados (sobre un millón), cadenas que tengan asociadas diferentes tipo de recompensas, la evaluación de medidas transientes y medidas estacionarias, etc.

Un aspecto práctico de nuestra investigación consiste en generar herramientas de software que permitan a los usuarios resolver problemas complejos de análisis de rendimiento por medio de programas fáciles de usar y que evalúen las métricas de



Reinaldo Vallejos, Sergio Yaksic, Jonathan Olavarría, Reinaldo Vallejos, José Manuel Martínez y Daniel Zuleta.

interés en muy poco tiempo (máximo unos pocos minutos). Para lograr la simplicidad en el uso de la herramienta se ha desarrollado una interfaz basada en grafos (que permite representar los modelos markovianos de los sistemas) y una interfaz de gráficos (para mostrar los resultados obtenidos). La interfaz de grafos permite representar grafos de miles de estados con atributos tales como scroll, 3D, zoom, tiempo real, entre otros. Debido a los buenos atributos de esta interfaz de grafos, como subproducto se han desarrollado otros tipos de aplicaciones como: representación de moléculas químicas, redes de computadores y sus algoritmos, y software para educación de matemática.

Alumnos

Se encuentran realizando su doctorado en esta área los estudiantes Daniel Zuleta y José Manuel Martínez. Los alumnos

encargados de desarrollar las herramientas de software son Sergio Yaksic y Jonathan Olavarría.

Publicaciones en Conferencias y Revistas

En los últimos tres años hemos publicado cuatro artículos en revistas ISI, un número similar en revistas no ISI, una veintena de artículos en congresos internacionales y un libro de texto.

Colaboraciones

Celso Ribeiro: UFF, RJ, Brasil; Marta Barría, UV, Chile; Alain Jean-Marie, INRIA-Montpellier, Francia; Gerardo Rubino, INRIA-Rennes, Francia; Héctor Cancela, Universidad de la República, Uruguay; Eduardo Moreno, Universidad Adolfo Ibáñez, Chile.



Equipo NIC Chile Research Labs.

NIC CHILE RESEARCH LABS

NIC Chile, Departamento de Ciencias de la Computación, Universidad de Chile.

NIC Chile Research Labs es un Laboratorio de Investigación Aplicada y Transferencia Tecnológica creado por NIC Chile. Su misión es desarrollar investigación de nivel internacional generando nuevos conocimientos en el área de redes IP, buscando siempre transferir sus resultados tanto a NIC Chile, como a la comunidad nacional y regional.

NIC Labs se focaliza principalmente en dos líneas de investigación y desarrollo: Convergencia de Tecnologías de Comunicación y Redes Avanzadas, incluyendo el futuro de Internet. Nuestro laboratorio tiene una política abierta a la comunidad: muchos de nuestros proyectos se encuentran a libre disposición para ésta.

Nuestro principal foco hasta este minuto ha sido establecer nexos entre NIC Research Labs, tanto con el gobierno como con la empresa privada. En este sentido hemos desarrollado un camino de entender, tomar, adaptar y aplicar investigación realizada en la academia y ponerla en la práctica desarrollando prototipos de software para nuestros clientes.

Dentro del desarrollo académico de nuestro laboratorio, contamos actualmente con posdoctorados realizando una de las más desarrolladas herramientas de skeletons para Java: Skandium. Este proyecto ha sido validado en varias publicaciones internacionales.

Además, como consecuencia del terremoto, hemos desarrollado un estudio del comportamiento de Internet durante este episodio, que ha sido presentado en diversas oportunidades e incluso publicado en la conferencia chilena de computación.

A continuación destacamos las publicaciones del laboratorio, tanto académicas como de divulgación a la comunidad.

Journals:

- Software: Practice and Experience

Conferencias:

- LNCS Euro-par
- IEEE Euro-micro PDP
- IEEE SCCC
- IEEE SCCC WSDP

Lista de Publicaciones:

1. Pablo Sepúlveda, Víctor Ramiro, Tomás Barros, José M. Piquer. Soundness of Chilean Networks. In XXIX International Conference of the Chilean Computer Society, 2010. (to appear).
2. Horacio González-Vélez and Mario Leyton. A Survey of Algorithmic Skeleton Frameworks: High-Level Structured Parallel Programming Enablers. In Software: Practice and Experience. (to appear).
3. Mario Leyton, Ludovic Henrio, and José M. Piquer. Exceptions for Algorithmic Skeletons. In LNCS Euro-par 2010.
4. Mario Leyton, José M. Piquer. Skandium: Multi-core Programming with Algorithmic Skeletons. In IEEE Euro-micro PDP 2010.
5. Mario Leyton, José M. Piquer. A Skandium based parallelization of DNSSEC. In IEEE SCCC WSDP 2009.

Otras Publicaciones y sitios de interés:

1. http://en.wikipedia.org/wiki/Algorithmic_skeleton
2. <http://skandium.niclabs.cl/>
3. <http://www.niclabs.cl/terremoto>

Software + Lenguajes

Alex Bergel:

CONSTRUIR CON CALIDAD

*Departamento de Ciencias de la Computación,
Universidad de Chile.*

Soy miembro del grupo PLEIAD del Departamento de Ciencias de la Computación de la Universidad de Chile.

Mi investigación se enfoca en ingeniería de software, particularmente en calidad de software. Las herramientas tradicionalmente utilizadas como lenguajes de programación y ambientes de desarrollo son tremendamente poderosas para construir software, pero son claramente insuficientes para todas las actividades de mantención. Es bien conocido que las empresas gastan un 75% de sus recursos en mantener software, en vez de producir otros nuevos. Mi actividad de investigación ofrece nuevas técnicas y metodologías para facilitar la mantención y el control de calidad de software.

Mis hipótesis de trabajo se basan en la utilización de herramientas de visualización y de metamodelización para ayudar a los desarrolladores a identificar deficiencias y anomalías en su propio código. Mis últimos resultados son Mondrian y Spy. Mondrian es una herramienta ágil para crear "mapas de software". Spy es un framework para construir perfiles de ejecución de código. Mondrian y Spy son parte de la plataforma de análisis de software Moose. Mondrian es un elemento central de Moose sobre el cual se utilizan la mayoría de las herramientas construidas con Moose.

Moose es co-desarrollado con INRIA Lille Nord-Europe (France), Universidad de Berna (Suiza), Universidad de Lugano (Suiza), Vrije Universiteit Brussel (Bélgica) y Universidad de Chile. En cada uno de estos sitios, ingenieros, doctorados e investigadores participan en un esfuerzo colectivo con una gran interacción (por



Alex Bergel, Romain Robbes, Felipe Bañados y Patricio Plaza.

ejemplo, co-escritura de artículos, feedback sobre herramientas, pair-programming, co-organización de eventos).

En la Universidad de Chile trabajo esencialmente con los profesores Cecilia Bastarica, Johan Fabry, Sergio Ochoa, Romain Robbes y Eric Tanter, y con los alumnos Felipe Bañados, Julio Hurtado, Christian Palomares y Vanessa Peña. Julio trabaja en la modernización de procesos de software. Su herramienta se llama Avispa y es desarrollada con Moose. Felipe trabaja en la diferenciación de perfiles de ejecución y es el autor de Hip, una extensión de Spy. Christian trabaja en el ambiente Seaside. Vanessa analiza la cobertura de los unit tests.

Publico los resultados de mi trabajo de investigación en las conferencias y revistas más competitivas. ECOOP, OOPSLA y TOOLS, son las conferencias que privilegio para difundir mis resultados académicos, ya que son reconocidas como las más prestigiosas en el área de programación con objetos. Mis revistas preferidas son Transaction on Software Engineering (TSE) y Elsevier Computer Languages, Systems and Structures. Mis herramientas de investigación son regularmente presentadas

a SPIN Chile (red de empresas que tienen un enfoque en la calidad de software), European Smalltalk User Group (ESUG) y Smalltalks.

Durante 2010, tuve ocho artículos aceptados en eventos internacionales (cuatro conferencias y cuatro workshops).

Cecilia Bastarica:

DESARROLLO DIRIGIDO POR MODELOS: UN NUEVO ENFOQUE EN INGENIERÍA DE SOFTWARE

*Departamento de Ciencias de la Computación,
Universidad de Chile.*

"La ingeniería de software es la aplicación de las Ciencias de la Computación para la resolución de problemas en presencia de recursos limitados."

El Departamento de Ciencias de la Computación (DCC) de la Universidad de Chile, está dentro de la Escuela de Ingeniería, y por lo tanto resulta natural



Cecilia Bastarrica.

reconocer que es importante darle un perfil ingenieril a la especialidad de Ingeniería Civil en Computación. Esto es aún más importante cuando gran parte de los alumnos titulados se desempeñan como ingenieros de software en el ámbito laboral. Sin embargo, el DCC había tenido tradicionalmente desde sus inicios y hasta alrededor de 1998, una orientación más científica que ingenieril.

Desde entonces se ha venido haciendo un esfuerzo sistemático por desarrollar la Ingeniería de Software dentro del Departamento. Es así como se contrataron varios profesores especialistas en ingeniería de software y se le ha dado, como consecuencia, una relevancia mayor a los cursos del área. También esto hace que tengamos el potencial de convertirnos en un polo poderoso en esta área tanto en Chile como en Latinoamérica.

Conceptualmente, la ingeniería de software tiene dos facetas: una más técnica y otra más relativa a la gestión de proyectos de software. Ambas han sido abordadas de manera conjunta en la docencia, pero han tomado rumbos independientes en lo relativo a la investigación y la transferencia tecnológica.

En 1998 se creó el Diploma en Gestión Informática, que luego evolucionó a lo que hoy imparte el DCC como Diploma de Postítulo en Gestión Informática. Fue el primer Postítulo de esa naturaleza en Chile, y aún hoy constituye uno de los programas más prestigiosos en su ámbito. Por su parte, la transferencia de los aspectos técnicos de la ingeniería de software se desarrolla

en el Postítulo en Ingeniería y Calidad del Software, creado en el año 2002, y que no tiene competencia en Chile hasta la fecha. Ambos postítulos constituyen la base de los cursos del Magíster en Tecnologías de la Información, que se imparte en el DCC y cuyos alumnos son esencialmente profesionales que buscan actualizar sus conocimientos luego de un tiempo de haberse titulado.

La investigación en Ingeniería de Software dentro del DCC ha tenido un desarrollo más lento para obtener sus mayores logros. Las primeras publicaciones en conferencias internacionales de alguna relevancia fueron en el área de ingeniería Web (LA-WEB, ICWE) o en la revista *Journal of Web Engineering*, entre los años 2002 y 2004. Sin embargo, esta orientación no ha prosperado, al menos en nuestro Departamento.

Más recientemente mayores logros se han obtenido en el área de diseño de software por parte del grupo MaTE creado en 2007. Claramente, contar con una masa crítica de investigadores le ha dado un gran impulso al área y le ha permitido publicar en las conferencias más relevantes del mundo. Esto se ha visto reflejado en que artículos generados en este grupo hayan sido elegidos dentro de la mejor investigación nacional en computación que se presentó durante las Jornadas Chilenas de Computación 2009, y también entre los mejores de 2010.

El grupo MaTE se ha especializado en desarrollo de software dirigido por modelos, y también ha publicado sus resultados en conferencias de primera línea internacional tales como ASE, WICSA, SPLC, ICSR, ICSP, SHARK e ICMT. Sus resultados también han sido publicados en revistas tales como *International Journal of Software Engineering and Knowledge Engineering*, *Journal of Software and Systems Modeling* y *Advances in Engineering Software*.

El grupo MaTE ha desarrollado un proyecto de investigación pura en colaboración con el INRIA, pero en general su labor ha estado más orientada hacia la investigación aplicada a la industria. Es en este contexto que se ha involucrado en el proyecto Tutelkán durante los últimos cinco años, que ha tenido como objetivo mejorar los estándares de desarrollo de software de la industria chilena. Este proyecto fue financiado por

CORFO y se realizó en colaboración con la Universidad Técnica Federico Santa María (UTFSM), la GECHS, la ACTI y SPIN-Chile. Recientemente también se adjudicó un nuevo proyecto Fondef de investigación aplicada y de interés público, conjuntamente con el mismo grupo de la UTFSM que participó de Tutelkán, referido a la formalización y adaptación automática de modelos de proceso de desarrollo de software.

Actualmente el grupo MaTE está formado por las académicas María Cecilia Bastarrica y Nancy Hitschfeld, colaborando regularmente con otros académicos del DCC tales como Sergio Ochoa y Alexandre Bergel. También colaboramos regularmente con los profesores de la UTFSM Marcello Visconti, Hernán Astudillo, Jocelyn Simmonds y Claudio Lobos, estos dos últimos ex alumnos del DCC. Hay cuatro estudiantes de Doctorado: Andrés Vignaga, Daniel Perovich, Pedro Rossel y Julio Ariel Hurtado, y tres estudiantes de Magíster: Alejandro Lagos, Christian Peña y Eduardo Sotomayor.

Contando en la actualidad con un gran grupo humano, muy competente, dedicado a la investigación en ingeniería de software y publicando al mejor nivel internacional, el DCC es sin duda el Departamento de computación más poderoso de Chile, en esta área. También sus múltiples proyectos han contribuido a que esta área sea actualmente la que concentra la mayor parte de los estudiantes de posgrado del DCC, procedentes en general de toda Latinoamérica.

Eric Tanter:

LENGUAJES DE PROGRAMACIÓN: HERRAMIENTAS FUNDAMENTALES PARA EL DESARROLLO DE SOFTWARE

Departamento de Ciencias de la Computación, Universidad de Chile.

Si no fuese porque los podemos programar, los computadores no serían tan fascinantes y versátiles. Pero, ¿cómo programar? En primer lugar, hay que tener un lenguaje para

comunicarse con la máquina. Expresarse en el lenguaje directamente comprendido por la máquina es extremadamente de bajo nivel (a nadie le gusta hablar sólo con 0s y 1s) y no permite aprehender programas de gran tamaño. Por eso, se han creado muchos lenguajes de programación llamados de “alto nivel”, y regularmente vemos nuevos lenguajes aparecer. ¿Por qué es así? ¿Por qué no estamos todos contentos y felices con el lenguaje X? (reemplace X por C, Java, C#, Python, Lisp, SQL, Ruby, Smalltalk, Scala, bash, Javascript, o cualquier otro de su gusto)

Cada lenguaje de programación refleja principios y objetivos considerados cruciales para los desarrolladores. Esos principios y objetivos pueden ser antónimos como lo son la eficiencia, la seguridad, y la flexibilidad. Además, cada vez que surge un dominio nuevo para programas (por ejemplo, la Web, los aparatos móviles, los controladores de cohetes), las necesidades son distintas, y los lenguajes existentes, que fueron concebidos con otros fines, se revelan inadecuados.

Mi investigación se centra en explorar distintas dimensiones de los lenguajes de programación para lograr proponer medios adecuados para el desarrollo de cualquier programa, por más complejo que sea. No se trata de buscar un lenguaje perfecto (desgraciadamente no existe), pero sí de buscar mecanismos específicos que pueden ayudar en ciertos casos. Me interesa proponer nuevas abstracciones, estudiar sus propiedades formales y las garantías que proveen, ver cómo se pueden implementar eficientemente, y proponer ambientes de programación que apoyen su uso. Más específicamente, trabajo en varios temas relacionados con la programación por objetos y por aspectos (una nueva abstracción que permite definir módulos que observan y controlan a otros). Me fascina la problemática de reconciliar la flexibilidad y adaptabilidad provistas por lenguajes dinámicos, con la necesidad de imponer barreras de abstracción y encapsulación, así como proveer ciertas garantías de seguridad.



Grupo PLEIAD: Eric Tanter, Johan Fabry, Rodolfo Toledo, Paul Leger, Guillaume Pothier, Ismael Figueroa, Óscar Callau y Esteban Allende.

Soy miembro co-fundador del laboratorio PLEIAD, que explora varias temáticas relacionadas con los lenguajes de programación y sus ambientes de programación. Además de trabajar con otros profesores de PLEIAD (Alex Bergel, Johan Fabry, Romain Robbes), estoy constantemente colaborando con investigadores fuera de Chile. He dirigido y participado en proyectos de investigación con universidades y centros de investigación en Estados Unidos, Canadá, Brasil, Francia, Bélgica, Holanda, Suiza, entre otros.

A la fecha, soy profesor guía de cinco estudiantes de Doctorado en la Universidad de Chile, y co-guía de un doctorando en Bélgica. Estos estudiantes, algunos cerca de titularse y otros recién empezando, trabajan en diversos temas como debugging, aspectos y seguridad, sistemas sensibles a su contexto de ejecución, sistemas de tipos graduales, sistemas distribuidos y concurrentes, y computación pervasiva.

En los últimos cinco años he publicado más de 50 artículos de investigación, incluyendo cerca de 20 artículos en revistas internacionales (IEEE Software, Science of Computer Programming, Software Practice & Experience, etc.). Publico regularmente en conferencias de prestigio en el área como ECOOP, OOPSLA, y AOSD. He participado en más de 30 comités de programa, incluyendo los de las conferencias

ECOOP y AOSD. Soy presidente del comité de programa de AOSD 2012, la conferencia de referencia en el desarrollo de software por aspectos.

Los lenguajes de programación son un área esencial de la computación, cubriendo un largo espectro desde lo teórico hasta lo muy aplicado. ¡Todo un programa!

Johan Fabry:

ASPECTOS DE DESARROLLO CON ASPECTOS

Departamento de Ciencias de la Computación, Universidad de Chile.

Un paradigma relativamente nuevo de programación, nacido de la Programación a Objetos, pero no limitado a ellos, es la Programación por Aspectos. La motivación por los aspectos, al igual que en muchas otras evoluciones en la historia de ingeniería de software, es obtener una mayor modularización del software. Donde objetos (y también otros paradigmas “clásicos”) fallan es en la modularización de funcionalidad (o preocupaciones) cuya implementación está esparcida en varias partes de la estructura de la aplicación, los llamados “cross-cutting



Johan Fabry (Fotografía: Comunicaciones FCFM).

concerns” (preocupaciones transversales). Aspectos son módulos que resuelven ese problema porque tienen no sólo la implementación de su comportamiento, sino además contienen la especificación de cuándo en la ejecución de los otros módulos del programa, esa funcionalidad debe ejecutarse (lo que se llama un *pointcut*). Así la funcionalidad esparcida se puede centralizar en un módulo.

Mi investigación se enfoca mayoritariamente en ayudar a los programadores que escriben código con aspectos. Entre otros, escribir código con aspectos requiere especificar los *pointcuts*, lo que puede resultar *non-trivial*. Mi trabajo reciente para ayudar en eso es *AspectMaps*: una visualización de dónde en el programa el aspecto ejecuta su funcionalidad. Eso permite ver, por un lado, si los *pointcuts* que uno escribe están correctos y, por otro, entender más fácilmente código ajeno (con aspectos). Otra parte de mi investigación es la combinación de lenguajes a dominio específico (DSL) y de los aspectos, resultando en lenguajes de aspecto a dominio específico: DSAL. Ese dominio de investigación busca juntar las ventajas conocidas de los DSL con la Programación por Aspectos. Destaco una ventaja específica aquí: DSLs hacen posible que gente *non experta* en programación pueda escribir programas. DSALs puede hacer posible que gente *non experta* en Programación con Aspectos pueda escribir programas con aspectos. Mi investigación en DSALs

aborda infraestructura y metodologías para crear estos lenguajes y también considera posibles interacciones entre varios aspectos escritos en varios lenguajes DSAL.

Junto a esto, también estoy interesado en otros paradigmas de modularización que podemos considerar como avanzados (por ejemplo, *Traits*) y otras herramientas de programación.

Como miembro co-fundador del laboratorio de investigación PLEIAD coopero mayoritariamente con otros profesores del laboratorio: Alex Bergel, Romain Robbes y Eric Tanter.

Con Alex desarrollamos la infraestructura de visualización utilizada por *AspectMaps*, y tenemos un proyecto *SticAmSud* con LIFIA (UNLP, Argentina) y RMoD (INRIA Lille Nord-Europe, Francia) donde mi parte del trabajo se centra en este tema.

En este proyecto, además del trabajo junto con los integrantes de RMoD, mi colaboración local es con Romain Robbes, con quien trabajamos en implementar una infraestructura de notificaciones para herramientas, de desarrollo. Tomando inspiración de conceptos de lenguajes de aspectos podemos ofrecer mayor soporte al creador de estas herramientas lo que facilita su creación.

Con Eric Tanter hemos publicado varios trabajos sobre infraestructura para el desarrollo de DSALs (trabajo llamado *ReLax*), y mayor expresividad para definir la aplicabilidad de aspectos en sistemas distribuidos (trabajo de *distributed scoping strategies*). Con Eric Tanter somos parte del *Equipe Associee INRIA* llamado *RAPIDS* que trata seguridad en sistemas distribuidos con aspectos, donde mi parte se enfoca en la creación de DSAL para varias facetas de distribución.

Aparte de estas cooperaciones, he trabajado junto a investigadores de la *Vrije Universiteit Brussel*, Bélgica, en el trabajo de *AspectMaps* y trabajos relacionados con DSALs.

Tengo dos estudiantes de Doctorado:

Esteban Allende: Su tema encaja con el trabajo de Óscar Callau, estudiante de

Doctorado de Eric Tanter. Esteban utilizará los tipos graduales dentro de la máquina virtual *Squeak* para realizar optimizaciones, logrando una mayor velocidad de ejecución de programas cuando tienen esa información de tipos.

Arturo Zambrano: estudiante de Doctorado del LIFIA, su profesor guía es Silvia Gordillo y yo soy profesor co-guía. El trabajo de Arturo consiste en una evaluación de las metodologías y herramientas de desarrollo con aspectos, en el ciclo completo de desarrollo de una aplicación. Nos enfocamos en la problemática de dependencias e interacciones, tomando un caso de estudio específico con el cual Arturo tiene amplia experiencia como desarrollador en la industria.

Conferencias claves para mi área de investigación son: *AOSD* (*Aspect-Oriented Software Development*), *ECOOP* (*European Conference on Object-Oriented Programming*), *ACM SAC Programming Languages and Programming for Separation of Concerns*. El journal ISI de preferencia es *Elsevier Science of Computer Programming*.

Romain Robbes:

INVESTIGACIÓN EN MINERÍA DE REPOSITARIOS DE SOFTWARE EN PLEIAD

Departamento de Ciencias de la Computación, Universidad de Chile.

Soy miembro del Grupo de Investigación PLEIAD. Mi investigación se centra en el tema de Ingeniería de Software, específicamente en el ámbito de la Minería de Repositorios de Software (*Mining Software Repositories* o MSR). La investigación en MSR explota la gran cantidad de datos producidos por los desarrolladores, probadores, mantenedores, etc. a fin de validar empíricamente la eficacia de varios enfoques que apoyan a estos profesionales durante el desempeño del trabajo relacionado con la evolución de los sistemas de software.



Romain Robbes.

Pero, ¿por qué los desarrolladores, probadores, mantenedores, y también los managers y arquitectos, necesitan este apoyo?

El mantenimiento y la evolución del software tiene un valor del 75% del costo total de un sistema de software; este porcentaje va en aumento debido a que los sistemas de software se utilizan durante extensos periodos de tiempo. El mantenimiento de software tiene un costo muy elevado por su complejidad. Incluso, realizar cambios simples, como por ejemplo cambiar el nombre de una función por uno más descriptivo, puede resultar un desafío si la función se utiliza miles de veces en una base de código de gran envergadura.

En este contexto, la investigación en MSR parte del supuesto de que la historia de un sistema de software contiene información extremadamente valiosa. Como dice Santayana: *“Aquellos que no pueden recordar el pasado están condenados a repetirlo”*. En el caso de sistemas de software, su historia se puede registrar con gran precisión, en lo que llamamos repositorios de software. Entre algunos repositorios de software encontramos: el sistema de control de versiones (por ejemplo, CVS, Subversion, Git), que contiene todos los cambios en el sistema, sus fechas, y los autores; el sistema de seguimiento de defectos (por ejemplo, Bugzilla, JIRA, Trac), que contiene todos los informes de problemas (bugs), que afectan

al sistema, su importancia y su estado; o archivos de listas de distribución de e-mails, que contienen todas las conversaciones sobre el sistema.

A continuación, dos ejemplos específicos de cómo MSR puede ayudar a los desarrolladores y managers en sus actividades diarias:

- **Predicción de los cambios:** Al observar la manera en que el sistema ha cambiado en el pasado, podemos inferir patrones de cambio. Por ejemplo: cuando un método “a()” cambia, hay una probabilidad del 90% de que el método “b()” cambie también. Si un desarrollador rompe el patrón –cambiando sólo “a()”– podemos aconsejarle que verifique si también es necesario cambiar “b()”. Si el desarrollador se ha olvidado efectivamente de cambiar “b()”, hemos impedido un error potencial.
- **Predicción de errores:** Cuando los recursos son limitados, un director de proyecto no puede permitirse probar todos y cada uno de los componentes del software del mismo modo. Un enfoque de predicción de errores le dirá cuáles son los componentes más propensos a tener defectos, para que así, pueda asignar más probadores para ellos.

Es posible encontrar en esta misma edición de BITS (ver página 2), una descripción más detallada de la investigación en MSR. Mi trabajo en esta área de investigación, en los últimos tres años, ha dado lugar a cuatro publicaciones en revistas (en las revistas Automated Software Engineering; Empirical Software Engineering; Science of Computer Programming; y Software Tools for Technology Transfer), catorce publicaciones en conferencias generales como ICSE y ASE, y conferencias especializadas como MSR, WCRE, ICPC, TOOLS y MoDELS.

Por supuesto, este trabajo no lo he hecho solo, sino que es fruto de colaboraciones con muchos otros investigadores.

En el pasado, he sido miembro del grupo de investigación REVEAL, en la Universidad de Lugano, en Suiza. He trabajado, y trabajo aún, con sus miembros: Prof. Michele Lanza, Dr. Mircea Lungu (ahora en el SCG

in Berna), Dr. Marco D’Ambros, Dr. Richard Wettel, Lile Hattori, Fernando Olivero, y Alberto Bacchelli.

Desde mi llegada al DCC de la Universidad de Chile, en enero de 2010, he colaborado activamente con otros miembros del grupo de investigación PLEIAD (profesores Alexandre Bergel, Johan Fabry y Éric Tanter), y otros profesores del DCC (Gonzalo Navarro y Sergio Ochoa). En este momento, no estoy trabajando con ningún estudiante, pero sí, estoy en la búsqueda. Así que si este breve relato de mi trabajo de investigación te resulta interesante, ¡no dudes en ponerte en contacto conmigo! (rrobbes@dcc.uchile.cl).

SIGSE: SPECIAL INTEREST GROUP ON SOFTWARE ENGINEERING

*Departamento de Ciencia de la Computación,
Pontificia Universidad Católica de Chile.*

La relación entre personas y computadores ha cambiado dramáticamente en los últimos años y esto se debe a que hoy en día los programas de computadora (o software) juegan un rol central en casi todos los aspectos de nuestra vida diaria, en el gobierno, bancos, finanzas, educación, transporte, entretenimiento, medicina, agricultura y leyes, por citar algunos ejemplos de aplicación.

Este crecimiento notable en la dependencia del uso de productos de software se debe a que éstos brindan a las personas, herramientas que les permiten ser más eficaces al resolver sus problemas y les proveen un medio para trabajar y entretenerse que es, a menudo, más seguro, más flexible y menos limitado que otros medios. Sin embargo, la naturaleza del software supone propiedades intrínsecas esenciales -complejidad, invisibilidad, flexibilidad, evolución- que son difíciles de abordar y que sumadas a requisitos actuales como la tendencia a crear productos de escala masiva, con tiempos de respuesta que se miden en segundos, que soportan diferentes esquemas de calidad tales como la seguridad,



Jaime Navón, Rosa Alarcón, Yadrán Eterovic y Andrés Neyem.

que pueden encontrarse distribuidos o requieren sofisticados mecanismos de coordinación, sitúan a los productos de software entre los sistemas más complejos hechos por el hombre.

El término *ingeniería de software* se usa hoy ampliamente en los sectores productivos y de servicios, en el gobierno y en las universidades. Ingeniería de software significa la aplicación de un enfoque sistemático, disciplinado y cuantificable de desarrollo, operación y mantenimiento de software. Se trata de algo más que simplemente producir código, incluye calidad, plazos y presupuestos, y el conocimiento y la aplicación de principios y disciplina. Se la puede definir como la ingeniería que aplica sistemáticamente y en forma disciplinada los principios de la Ciencia de la Computación y las matemáticas para lograr soluciones confiables y económicas a problemas de software, así como para la operación y el mantenimiento de estas soluciones. La ingeniería de software también es diferente en carácter a otras disciplinas de la ingeniería, debido tanto a la naturaleza intangible del software como a la naturaleza discreta de la operación del software.

El Departamento de Ciencia de la Computación (DCC) de la Pontificia Universidad Católica de Chile ha estado involucrado en ingeniería de software por más de una década. Un ejemplo de esto es que hemos sido pioneros en incorporar el lenguaje de modelado UML

en el programa de estudios de pregrado y también en apoyar su introducción en la industria a nivel nacional realizando asesorías y programas de capacitación en empresas grandes y medianas. Esto último se inserta también en la preocupación del DCC por la transferencia tecnológica. Otro ejemplo de este interés fue la ejecución de un proyecto Fondef que buscaba desarrollar frameworks de aplicación (patrones más componentes de software) para la banca cuando el término framework aún no era conocido por el ambiente local.

En cuanto a la formación de profesionales, a partir de 2009, los alumnos que ingresan a computación pueden elegir la especialidad de Ingeniería de Software. Esta especialidad se enfoca en la formación de ingenieros capaces de diseñar, implementar, validar, operar y mantener sistemas de software como soluciones a problemas reales, satisfaciendo las necesidades de los clientes y usuarios y las restricciones presupuestarias y de tiempo. El programa curricular, de cinco años, proporciona una sólida formación inicial común en ciencias básicas, ciencias de la ingeniería, e ingeniería industrial, y luego se concentra en la formación especializada: Ciencia de la Computación, e ingeniería de software propiamente tal. En esta última área se estudia los conceptos y se ponen en práctica las técnicas y herramientas asociados a la gestión y ejecución de proyectos, la arquitectura de sistemas y el diseño detallado de software, y el testing. El propósito es

brindar al profesional conocimiento integral y experiencia práctica para el desarrollo de sistemas en ambientes complejos.

Desarrollando investigación en ingeniería de software

En el año 2008, nace SIGSE (Special Interest Group on Software Engineering), un grupo de investigación en ingeniería de software, actualmente formado por cuatro profesores que colaboran en la supervisión del trabajo de varios alumnos de doctorado, alumnos de magíster y alumnos memoristas. Una de las premisas de SIGSE es generar conocimiento de relevancia nacional e internacional. La investigación que realiza nuestro grupo abarca diversas áreas sobre problemas relacionados con procesos de desarrollo, diseño y arquitectura de software, arquitecturas orientadas a servicios y computación móvil. A modo de ejemplo, detallamos algunas de estas líneas de investigación:

(1) Computación Orientada a Servicio:

La Computación Orientada a Servicios (SOC - Service Oriented Computing) es un paradigma que ha ganado mucha atención en la industria del software debido a que representa una nueva forma de desarrollar arquitecturas de sistemas distribuidos. SOC es una evolución de la ingeniería de software basada en componentes que introduce un nuevo tipo de bloque de construcción llamado servicio, el cual es una funcionalidad que es consumida remotamente utilizando protocolos estándares. A pesar de los importantes beneficios que proporciona este paradigma, aún siguen pendientes temas relacionados con decrecer los costos de creación y mantenimiento de este tipo de aplicaciones. Por ejemplo, los desarrolladores tienen que invertir bastante esfuerzo en descubrir los servicios manualmente, proporcionar el código para invocarlos y realizar las modificaciones necesarias durante la fase de mantenimiento. Los servicios pueden además invocarse entre ellos dando

lugar a servicios compuestos, cuyas partes proveen diferentes niveles de calidad y podrían no estar disponibles en tiempo de ejecución. Una de las líneas de investigación que nuestro grupo realiza es proporcionar mejoras a este paradigma mediante la incorporación de ontologías, metadatos semánticos, y técnicas de razonamiento que permitan la composición dinámica de servicios y la garantía de diferentes niveles de calidad. La investigación se enfoca en el diseño y desarrollo de plataformas para crear servicios Web semánticos así como en la incorporación de enfoques ligeros como REST y técnicas de la Web 2.0 tales como Mashups.

(2) Diseño y arquitectura de software Web:

El DCC ha sido pionero también en incorporar las tecnologías de la Web a sus programas de estudio a nivel de pregrado y posgrado. Rápidamente las páginas Web se transformaron en sitios para finalmente aparecer como aplicaciones Web. El desarrollo para la Web tiene características especiales tanto en el proceso como en la arquitectura de la solución. En el proceso mismo, la participación de diseñadores y el diálogo de éstos con los ingenieros ha sido una problemática que ha motivado incluso el surgimiento de nuevas tecnologías. El patrón modelo-vista-controlador ya conocido cobra una nueva relevancia y dimensión, y aparecen los diversos frameworks que permiten facilitar una tarea que al principio era bastante dura. En la actualidad se llevan a cabo trabajos y tesis de alumnos en temas tan diversos como arquitecturas para customización en el lado del cliente que permitan mantener las ventajas de ubicuidad de la aplicación Web; arquitecturas para aplicaciones autoadaptables de acuerdo a frecuencia de navegación de los usuarios; sintetización de un servicio Web en forma automática a partir de un sitio Web, etc.

(3) Procesos de desarrollo:

El desarrollo de software ha sido, desde sus inicios históricos, hace más de 50 años, una tarea difícil. Los proyectos tienden a exceder sus plazos y presupuestos, y a ofrecer

menos funcionalidad y de peor calidad que la inicialmente acordada. Para mejorar esta situación, investigadores y profesionales han propuesto modelos de procesos de desarrollo de software basados en las prácticas que aplican habitualmente las organizaciones que son exitosas desarrollando software. Sin embargo, la adopción de un proceso de desarrollo o, incluso, de algunas de sus prácticas, tampoco es una tarea simple. Esta línea de investigación aplicada busca, por una parte, determinar el proceso más adecuado a las características de un proyecto u organización particular y, por otra, formas eficaces de llevar adelante la adopción de prácticas tales como gestión de requisitos, estimaciones de esfuerzo y plazos, gestión de riesgos, modelado visual, y desarrollo iterativo e incremental. Para esto, se debe realizar una serie de actividades, tales como convencer a algunas personas y entrenar a otras, ayudar a elegir el proyecto adecuado para iniciar la aplicación del nuevo proceso de desarrollo y supervisar la ejecución del mismo por el equipo de desarrollo, identificando fortalezas y debilidades. En resumen, nuestro objetivo es ayudar a las organizaciones de desarrollo de software a ser mejores en lo que hacen.

(4) Computación móvil:

En las últimas tres décadas, nuestra relación con la computación ha estado centrada principalmente en el PC e Internet. Hoy, los avances en computación móvil y comunicaciones inalámbricas, nos están llevando hacia una nueva relación con la tecnología caracterizada principalmente por una fuerte inclusión de la computación en las actividades diarias llevadas a cabo por las personas (por ejemplo en el ocio y entretenimiento, en operaciones financieras y en la educación, entre otros). El desarrollo de software para dispositivos móviles es una tarea que impone al desarrollador lidiar con nuevo desafíos originados por la diversidad de hardware y sensores, sistemas operativos, plataformas de desarrollo y escenarios de aplicación. Esta línea de investigación se orienta a brindar soluciones a los desarrolladores para facilitar la creación de

aplicaciones móviles, a través de frameworks de desarrollo que encapsulan estrategias de diseño que abordan problemas recurrentes de un dominio específico de aplicación.

Resumiendo, el DCC en general y SIGSE, en particular, buscan contribuir al mejoramiento del marco conceptual y de la práctica de la ingeniería de software, tanto en Chile como internacionalmente. Para esto, desarrollamos actividades de investigación, de formación de capital humano de pre y posgrado, y de transferencia tecnológica. Como resultado, esperamos transformarnos en el mediano plazo en un referente latinoamericano en el área.

Journals

- IEEE Software
- Computers and Education
- Computing and Informatics
- Expert Systems with Applications
- Journal of Group Decision and Negotiation

Conferencias

- WWW (World Wide Web)
- ECTEL (European Conference on Technology Enhanced Learning)
- IASTED (International Conference on Software Engineering and Applications)
- SEDE (International Conference on Software Engineering and Data Engineering)
- IWSSA (Workshop on System/Software Architectures)
- CSCWD (Computer Supported Cooperative Work in Design)
- CRIWG (Workshop on Groupware: Design, Implementation, and Use)

Más información sobre nuestro grupo puede ser obtenida en sigse.ing.puc.cl

Algoritmos + Criptografía + Estructura de Datos

CLCERT: CRIPTOGRAFÍA APLICADA Y SEGURIDAD

Departamento de Ciencias de la Computación, Universidad de Chile.

El CLCERT, grupo de criptografía aplicada y seguridad de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile, focaliza su investigación en dos áreas. La primera es Criptografía Aplicada, incluyendo el diseño y estudio de protocolos orientados a mejorar la privacidad de los participantes; sistemas de comunicación anónima segura; votación electrónica verificable, y en general, computación distribuida segura. La segunda es Seguridad Aplicada, principalmente en el estudio de phishing y malware.

El grupo está compuesto por el profesor Marcos Kiwi (Departamento de Ingeniería Matemática, Universidad de Chile), el ingeniero Sergio Miranda (Universidad de Chile); los estudiantes de Doctorado Philippe Camacho, y Julio Quinteros; los estudiantes de Magíster Gaston L'Huillier, Patricio Seguel; y los actuales estudiantes memoristas Alonso González, Renata Faccilongo, Rodrigo Porras, Francisca Merino, y Felipe Troncoso. El director del grupo es el profesor Alejandro Hevia (Departamento de Ciencias de la Computación, Universidad de Chile).

En términos de investigación en las áreas mencionadas, nuestro grupo colabora con investigadores como Tamara Rezk (INRIA Sophia Antipolis, Francia) y Alfredo Viola (Universidad de la República, Uruguay), aunque a la lista de coautores se ha incorporado recientemente a Gilles Barthe (IMDEA Software Institute, Madrid), Bogdan Warinschi (University of Bristol, Inglaterra), Richard Weber y Sebastián A. Ríos (Departamento de Ingeniería Industrial, Universidad Chile), y Daniele Micciancio (Universidad de California San Diego,



Sergio Miranda, Alejandro Hevia y Marcos Kiwi.

EE.UU.). Las publicaciones del grupo en conferencias y/o seminarios incluyen:

- IEEE Computer Security Foundations Symposium (CSF, IEEE Computer Society).
- Intl. Conference on Cryptology and Information Security (Latincrypt, LNCS Springer).
- IEEE International Conference on Intelligence and Security Informatics (ISI, IEEE Press).
- Intl. Conference on Information Security (ISC, LNCS Springer).
- Privacy Enhancing Technologies (PETS, LNCS Springer).
- Intl. Conference on Theory and Applications of Cryptographic Techniques (Eurocrypt, LNCS Springer).

En cuanto a publicaciones en revistas, éstas incluyen artículos en:

- Theoretical Computer Science (TCS, Elsevier).
- IEEE Selected Areas in Communication (JSAC, IEEE Press).

- Journal Computer and System Sciences (JCSS).

ACM Transactions on Information System Security (TISSEC, ACM Press).

Gonzalo Navarro:

ALGORÍTMICA Y TEORÍA DE LA INFORMACIÓN

Departamento de Ciencias de la Computación, Universidad de Chile.

La mayor parte de mi investigación reciente se enfoca en la intersección del área de Algoritmos y Estructuras de Datos, y la Teoría de la Información. El objetivo es desarrollar estructuras de datos que utilicen poca memoria y alcancen una eficiencia comparable a la de las estructuras clásicas. Esto tiene interés por la diferencia creciente entre el desempeño de los distintos niveles de la jerarquía de memoria, y la disponibilidad de memorias mayores en todos esos niveles. El utilizar menos espacio hace que una

estructura de datos pueda residir en una memoria más pequeña y rápida, con lo cual puede resultar que, a pesar de requerir más operaciones que su versión clásica, la estructura compacta resulte ser más rápida. En particular, cuando la estructura compacta cabe en memoria RAM y la clásica necesita utilizar el disco, la diferencia de desempeño puede ser de varios órdenes de magnitud. Estas estructuras también son de interés en dispositivos de capacidad limitada, como celulares, sensores, routers, etc.

Muchos de los desarrollos se concentran en el área de bases de datos de texto, donde se han conseguido avances espectaculares en la última década, tales como los llamados auto-índices. Estos representan un texto en un espacio cercano a su versión comprimida, pero dentro de ese espacio ofrecen búsqueda indexada, es decir de tiempo sublineal en el largo del texto. Las aplicaciones de este tipo de índices en áreas que necesitan manejar colecciones gigantescas de secuencias, como la bioinformática, recién están comenzando a explorarse. También hay varios resultados importantes para bases de datos de textos en lenguaje natural y en recuperación de información, donde la Web es un ejemplo obvio de la necesidad de utilizar el espacio en forma eficiente.

Como suele ocurrir, la investigación en esas áreas de aplicación ha llevado progresivamente a identificar problemas más básicos, donde se requieren estructuras de datos compactas para manejar árboles, secuencias, grillas, sumas parciales, grafos, relaciones binarias, permutaciones, y muchas otras. A su vez, los nuevos resultados en estructuras de datos básicas han dado lugar a resultados aplicados a problemas como indexación comprimida de colecciones XML, de objetos geográficos, de grafos Web y redes sociales, de índices invertidos, árboles de sufijos, y otras estructuras de relevancia para áreas como bioinformática, recuperación de información, sistemas de información geográficos, máquinas de búsqueda Web, etc.

Investigo también en otras áreas como Búsqueda por Similitud, Búsqueda Secuencial

e Indexada en Texto, y Algoritmos y Estructuras de Datos en general. Con respecto a la primera área cabe destacar la creación en 2008 de la conferencia SISAP (Similarity Search and Applications), para focalizar la investigación que se realizaba en los aspectos algorítmicos de la búsqueda por similitud.

Software

Estoy convencido de la importancia de prestar atención tanto al componente teórico como al práctico en la investigación en computación. Si falta lo primero se cae en heurísticas sin fundamento teórico, que no se comprende por qué funcionan ni en qué contexto dejarán de funcionar, y se alejan de la ciencia, donde es fundamental comprender los porqués. Si falta el componente práctico, al menos en computación se cae fácilmente en investigación teórica sin ninguna conexión con la realidad ni probable utilidad.

En algoritmos, esto se traduce en que debe haber un componente de diseño, uno de análisis teórico y uno de experimentación. Intento además que los desarrollos experimentales se conviertan en prototipos públicamente disponibles, para que sean usados con fines de investigación, docencia, y a veces incluso comerciales. Como consecuencia, hay bastante software públicamente disponible en <http://www.dcc.uchile.cl/gnavarro/software>, y otros sitios mucho más sofisticados desarrollados por mí o por mis alumnos, tales como <http://pizzachili.dcc.uchile.cl> y <http://www.recoded.cl>. No creo en la utilidad de las patentes en computación, y en general estoy en contra de patentar ideas abstractas tales como algoritmos.

Colegas y alumnos

Trabajo mucho mejor con un buen partner que solo. Como resultado, casi todas mis publicaciones son con coautores y tengo una larga lista de colaboradores en todo el mundo. Un listado de muestra, considerando sólo los más recientes y recurrentes, y tomando



Gonzalo Navarro.

sólo los jefes de grupos de investigación, incluye a Jérémy Barbay (Universidad de Chile), Nieves Brisaboa (Universidad de la Coruña, España), Edgar Chávez (Universidad Michoacana, México), Paolo Ferragina (Università di Pisa, Italia), Johannes Fischer (KIT, Alemania), Veli Makinen (Universidad de Helsinki, Finlandia), Sebastián Maneth (NICTA, Australia), Simon Puglisi (RMIT, Australia), Luís Russo (Universidad Nova de Lisboa, Portugal), y Kunihiro Sadakane (Universidad de Tokyo, Japón).

A esta lista debo agregar a mis ex-alumnos y ex-posdocs, con la mayoría de los cuales mantengo una relación de colaboración: Joaquín Adiego (PhD, U. Valladolid, España), Diego Arroyuelo (PhD, Yahoo! Research Chile), Benjamin Bustos (MSc, U. de Chile), Rodrigo Cánovas (MSc, U. de Chile), Francisco Claude (MSc, doctorando U. Waterloo, Canadá), Antonio Fariña (PhD, U. Coruña, España), Karina Figueroa (PhD, U. Michoacana, México), Travis Gagie (posdoc, U. Aalto, Finlandia), Rodrigo González (PhD, Index Technologies, U. de Chile), Gilberto Gutiérrez (PhD, U. Bío-Bío), Rodrigo Paredes (MSc y PhD, U. de Talca), y Diego Seco (posdoc, U. Coruña, España).

Finalmente, la mayoría de mis alumnos y posdocs actuales también son importantes colaboradores para la investigación: Carlos Bedregal (PhD), Ana Cerdeira (PhD, U. Coruña, España), Violeta Chang (PhD),

Cecilia Hernández (PhD), Norma Herrera (PhD, U. San Luis, Argentina), Sebastián Kreft (MSc), Fernando Krell (MSc), Susana Ladra (PhD, U. Coruña, España), Miguel Ángel Martínez (posdoc, U. Valladolid, España), Eliana Providel (MSc), Nora Reyes (PhD, U. San Luis, Argentina), Carina Ruano (MSc, U. San Luis, Argentina), y Daniel Valenzuela (MSc).

Busco completar la formación de mis alumnos enviándolos a congresos internacionales, en particular para que presenten los artículos en los que participan, e incluso a veces como oyentes si se lo han ganado con la calidad de su trabajo. Organizo siempre que puedo los “Miércoles de Algoritmos”, reuniones donde se exponen nuevas ideas, se analizan artículos de la literatura, se practican charlas, etc.

Publicaciones

Una lista de las conferencias relevantes donde he publicado más recientemente (últimos cinco años) incluye: ACM-SIAM SODA, STACS, ESA, ICDE, ACM SIGIR, CPM, ISAAC, DCC, LATIN, SPIRE, SEA, ACM-SIAM ALENEX, ACM RECOMB, y MFCS.

Una lista similar de revistas incluye: ACM Trans. Alg. (TALG), ACM Trans. Inf. Sys. (TOIS), ACM Trans. Web (TWEB), ACM Comp. Surv. (CSur), ACM J. of Exp. Alg. (JEA), Algorithmica, Theor. Comp. Sci. (TCS), Softw. Pract. Exp. (SPE), J. Comp. Biol. (JCB), Inf. & Comp. (IC), e Inf. Retr. (IR).

Desde 2006 he publicado 35 artículos en revistas internacionales y 62 en conferencias internacionales.

Otras distinciones

He sido Chair del Comité de Programa de siete congresos internacionales, estoy en el Comité Editorial de dos revistas internacionales (ACM JEA e IR), y he dado charlas plenarios en ocho congresos internacionales, entre otras muchas de menor importancia. En 2008 recibí el Premio Scopus, de Elsevier y Conicyt, en el área de Matemáticas, Computación e Ingeniería, a los autores más prolíficos de Chile.



Jérémý Barbay junto a alumnos del curso Alice.

Jérémý Barbay:

ANÁLISIS ADAPTATIVO: MÁS PRECISO, RÁPIDO Y PEQUEÑO

*Departamento de Ciencias de la Computación,
Universidad de Chile.*

Mi nombre es Jérémý Barbay. Nací y estudié en Francia, trabajé seis años en Canadá, y llevo tres años en Chile. Soy matemático por formación, teórico en Ciencias de la Computación por vocación, y usuario de computadores por hobby. Mi tema principal de investigación se relaciona con refinar las técnicas de análisis de “rendimiento de los algoritmos” y del “espacio de las estructuras de datos”. Esto, dentro de otros intereses como el mejoramiento de las técnicas de docencia, la teoría de la evolución y el diseño de mecanismos sociales en la red.

Mi ejemplo favorito de un problema práctico que requiere un análisis más fino de complejidad que el tradicional es la “intersección de arreglos ordenados”, que ocurren, por ejemplo, cuando los motores de búsqueda como Google tratan de resolver las consultas de sus usuarios. Dadas tres palabras u,v,w , correspondientes a tres arreglos ordenados U,V,W con las

referencias a las páginas asociadas con u,v,w , respectivamente, se pide responder la consulta “ u,v,w ”; es decir, se requiere buscar las referencias que los tres arreglos U,V,W tienen en común.

Un análisis tradicional agrupa las instancias por tamaño, y, además, analiza el comportamiento de los algoritmos en términos de su comportamiento “en el peor caso”. Este análisis intenta encontrar cuál es la peor clase de consulta que podría ser dada como entrada al algoritmo. Por ejemplo, en el caso del problema mencionado antes, una instancia de entrada que es muy compleja de tratar por cualquier algoritmo “razonable” es la siguiente: $U=\{1,2, 4,5, 7,8\}$, $V=\{1, 3,4, 6,7, 9\}$ y $W=\{ 2,3, 5,6, 8,9\}$. De hecho, para validar su resultado sobre esta instancia, cualquier algoritmo de intersección tiene que indicar nueve comparaciones, esto es, el tamaño del conjunto.

Sin embargo, tales instancias son muy artificiales, y, además, bastante diferentes a las que ocurren realmente en la práctica. El problema con el análisis de complejidad “en el peor caso” es, por tanto, que no logra diferenciar los distintos tipos de algoritmos con respecto a su “performance” sobre instancias que efectivamente ocurren en la práctica. Esto se debe a que las instancias prácticas son más fáciles y el análisis

tradicional es demasiado pesimista. Por otro lado, un análisis más fino de complejidad identificaría un parámetro adicional como, por ejemplo, la cantidad de comparaciones que se necesita por demostrar el resultado. Este medida es muy alta para el ejemplo que mostramos arriba, pero es mucho más baja para una instancia como: $U=\{1,2,3,7,8,9\}$, $V=\{4,5,6,7,8,9\}$ y $W=\{1,2,3,4,5,6\}$, que ocurre más usualmente en la práctica. Es por eso que un análisis fino puede ser un mejor predictor de la performance práctica de un algoritmo.

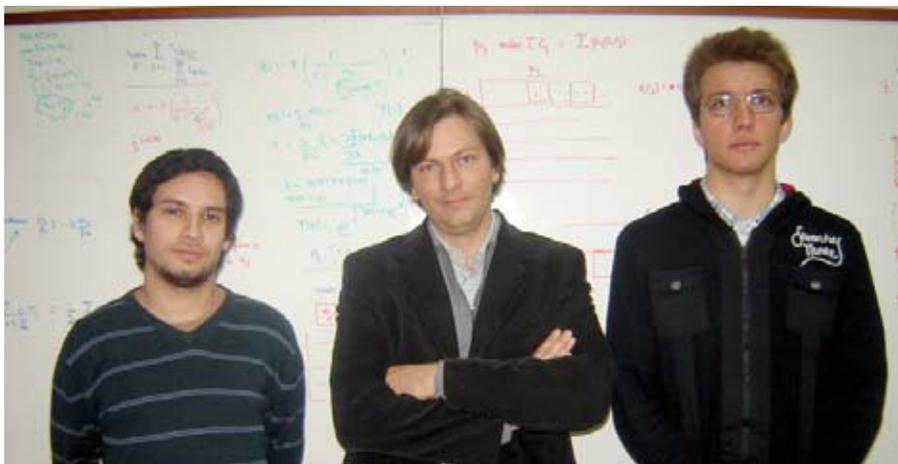
El problema de intersección es solamente un ejemplo, ya que también trabajo en variantes del problema de intersección, en algoritmos de ordenamiento, en problemas de geometría computacional, y en el análisis y diseño de estructuras de datos comprimidas, en colaboración con Gonzalo Navarro y Carlos Bedregal. Estos resultados fueron presentados en las conferencias de más alto nivel del campo, como SODA y FOCS, a razón de una o dos al año, y publiqué algunos en revistas prestigiosas como "ACM Transaction of Algorithms" y "Algorithmica". Enseño estas técnicas en cursos avanzados y también en cursos básicos de Ciencia de la Computación: el objetivo es enseñar a los alumnos este tipo de análisis lo más temprano posible.

José Rafael Correa:

ENTRE INVESTIGACIÓN DE OPERACIONES Y TEORÍA DE LA COMPUTACIÓN

Departamento de Ingeniería Industrial, Universidad de Chile.

Tras estudiar Ingeniería Matemática en la Universidad de Chile, partí a hacer un Doctorado en Investigación de Operaciones al MIT. En este marco me dediqué a trabajar en el diseño y análisis de algoritmos de aproximación para problemas de optimización combinatorial NP-difíciles. Lo que se busca en este contexto son algoritmos eficientes (a tiempo polinomial) que entreguen soluciones con una garantía de aproximación. Así pues, la pregunta



Omar Larré, José Rafael Correa y Charles Thraves.

fundamental que está detrás es: ¿Qué podemos hacer con un problema si restringimos el tiempo de ejecución a ser polinomial?

De regreso en Chile desarrollé también un interés por el estudio de los algoritmos "en línea" donde la idea es diseñar algoritmos que obtengan buenas soluciones, a pesar de que el input no es conocido de antemano sino que se revela en el tiempo. La pregunta en este caso es: ¿Qué podemos hacer con un problema si restringimos la información disponible?

Finalmente, en los últimos años me he interesado cada vez más en la Teoría Algorítmica de Juegos. Esta área, que ha visto muchos desarrollos en la última década, estudia sistemas distribuidos, en que diversos agentes toman decisiones en forma simultánea. Un tema que me ha interesado particularmente en este ámbito es el estudio del llamado "Precio de la Anarquía", el cual cuantifica la pérdida de optimalidad de un sistema descentralizado respecto de una solución coordinada. En este contexto, el paradigma de la eficiencia computacional, en el caso de los algoritmos de aproximación, o el paradigma de información de los algoritmos en línea se reemplazan por el paradigma de la coordinación entre los distintos agentes del sistema.

En este último tema trabajo actualmente con dos alumnos del Magíster de Gestión de Operaciones del Departamento de Ingeniería Industrial de la Universidad de Chile: Omar Larré y Charles Thraves. Omar

estudia el problema de ruteo en un grafo donde múltiples agentes buscan llegar a su destino en el menor tiempo posible. La dificultad es que los links sufren congestión, por lo que cuando muchos agentes quieren usar un determinado arco del grafo, los tiempos de todos los usuarios de ese arco se ven afectados en forma negativa. Por otra parte, Charles trabaja en un problema de políticas de precio con consumidores estratégicos. En este contexto, una firma quiere desarrollar un plan para vender un determinado producto, en que el precio irá decreciendo en el tiempo (liquidación). Los consumidores, entonces, se enfrentan a la disyuntiva de comprar hoy a precio normal o esperar a que éste baje. Pero esperar puede significar que el producto ya no esté disponible.

Como se desprende de lo anterior, mi área de investigación está en la frontera entre Investigación de Operaciones y Teoría de la Computación. Así pues, con frecuencia participo en reuniones de ambas comunidades. En febrero de 2010, por ejemplo, asistí a un workshop de computación en Dagstuhl, Alemania, mientras que en septiembre di una charla en el Departamento de Gestión de Operaciones de la Escuela de Negocios de la Universidad de Nueva York.

Lo mismo ocurre con las publicaciones. He publicado en conferencias de computación teórica como ICALP, IPCO, SODA, STOC y WINE, así como en revistas de Investigación de Operaciones como Operations Research y Mathematics of Operations Research.

Sistemas de Información



Ma. Andrea Rodríguez.

Ma. Andrea Rodríguez:

INVESTIGACIÓN EN SISTEMAS DE INFORMACIÓN ESPACIAL EN DIICC-UDEC

Departamento de Ingeniería Informática y Ciencias de la Computación, Universidad de Concepción.

Desde mis estudios de posgrado mi investigación ha estado centrada en el manejo de información geográfica (espacial) y espacio-temporal. Sistemas tales como monitoreo ambiental o planificación territorial, buscadores y servidores de mapas en la Web (Google Earth, Google Map), localización automática de vehículos y sistemas de navegación o ruteo, entre otros, son las aplicaciones inmediatas de la investigación que llevo a cabo. Estos sistemas representan objetos que son localizados en un espacio de más de una dimensión y en un instante o intervalo de tiempo. La información espacial es compleja, ya que debe lograr representar la geometría de los objetos y satisfacer las restricciones impuestas por el dominio espacial bajo representación. Así mismo, estos datos deben ser manipulados

por un conjunto de operadores que definen características, tales como área, perímetro o largo, y definen relaciones espaciales, tales como adyacencia, inclusión o separación. El manejo de este tipo de información requiere del desarrollo de modelos conceptuales y lógicos, estructuras de datos y algoritmos de procesamiento de información espacial y espacio/temporal.

Mi trabajo de investigación ha estado apoyado por el financiamiento obtenido desde Fondecyt, Centro de Investigación de la Web, ECOS/CONICYT y Fundación Andes. Las temáticas que se han cubierto abarcan distintos aspectos en el manejo de información espacial. A un nivel semántico se propuso funciones de similitud entre conceptos espaciales definidos en una o varias ontologías. A un nivel de estructuras de datos se propuso un meta-índice para resolver consultas espacio-temporales en un ambiente de servidores de datos distribuidos. En el ámbito de buscadores en la Web se implementó una estrategia de georeferenciación de documentos Web y de agrupamiento de noticias con referencia espacial. Mi trabajo más reciente ha abordado la formalización de restricciones de integridad espacial y espacio-temporal y el manejo de inconsistencias espaciales. El estudio considera modelos de bases de datos espaciales que son extensiones al modelo relacional para los cuales define restricciones de integridad que combinan atributos temáticos y geométricos. En una primera etapa fueron consideradas restricciones para información espacial estática y actualmente se están estudiando restricciones de integridad para regiones que evolucionan en el tiempo. Esta formalización ha permitido luego analizar semánticas de reparación de bases de datos inconsistentes y definir medidas de consistencia para caracterizar una base de datos inconsistente. El producto de mi trabajo de investigación se encuentra en publicaciones en revistas, tales como: IEEE Transactions on Knowledge and Data Engineering, Information Systems, IEEE Evolutionary Computation, International

Journal of Geographic Information Science, y en conferencias internacionales, tales como: Symposium on Spatial and Temporal Databases SSTD, ACM SIGSPATIAL GIS, Database Systems for Advanced Applications DAFTA, Web Information Systems Engineering WISE, entre otras.

En mi investigación he contado con colaboración a nivel nacional e internacional. En forma cercana he trabajado con Loreto Bravo (Universidad de Concepción) en la formalización de restricciones de integridad, y con Mónica Caniupán (Universidad del Bío-Bío) y Leopoldo Berstossi (University of Carleton, Canadá) en la definición de una semántica de reparación de bases de datos espaciales inconsistentes. Con Mauricio Marín (Yahoo Research! y Universidad de Santiago) he mantenido una constante colaboración en el trabajo de estructuras de datos para objetos en movimiento, lo que se basa en un trabajo previo con Gonzalo Navarro (DCC Universidad de Chile) y con el entonces alumno de doctorado del DCC Gilberto Gutiérrez (Universidad del BíoBío). Junto a Claudio Gutiérrez (DCC Universidad de Chile) he explorado propiedades topológicas de redes y comparto el interés por aplicar conceptos de Web semántica en el contexto de Linked Data y, en particular, Geo-Linked Data. A nivel internacional he trabajado, entre otros, con Max Egenhofer (supervisor de mi tesis doctoral) y con Fred Fonseca (Penn State University) en aspectos de ontologías para información espacial. Actualmente mantengo investigación conjunta con Nieves Brisaboa (Universidad de A Coruña) en cuanto a medidas de inconsistencia y con Christophe Claramunt (Naval Research Institute, France) para la modelación de restricciones de integridad espacio-temporal. No menos importante ha sido la colaboración de estudiantes de pregrado y del Magíster en Ciencias de la Computación de la Universidad de Concepción. En estas temáticas se han graduado diez alumnos de Magíster y más de quince alumnos de la carrera de Ingeniería Civil Informática de esta Universidad.

Benjamin Bustos:

CONTENT-BASED MULTIMEDIA INFORMATION RETRIEVAL

*Departamento de Ciencias de la Computación,
Universidad de Chile.*

Mis principales áreas de investigación se centran en las áreas de búsqueda por similitud en colecciones de datos multimedia, especialmente colecciones de imágenes en la Web, modelos 3D y secuencias de video, y en el área de algoritmos de indexamiento para información no estructurada, con énfasis en el manejo de colecciones muy grandes de información multimedia.

En particular, he desarrollado algoritmos y técnicas de indexamiento para espacios métricos, no métricos y multimétricos.

Algunos proyectos de investigación recientes en los cuáles he participado son los siguientes:

- (2010) Investigador (contraparte chilena) del Proyecto SCHR 1229/2-1 "German-Chile Research Cooperation on 3D Object Retrieval", financiado por la Fundación Alemana de Ciencia (DFG) dentro del Programa de Cooperación Chileno-Alemana en Investigación.
- (2007-2009) Investigador principal del Proyecto FONDECYT 11070037, "Effective and efficient retrieval in multimedia databases".
- (2007-2008) Investigador Joven en el Núcleo Milenio Centro de Investigación de la Web.

Colaboradores internacionales y nacionales

Colaboro con investigadores nacionales e internacionales en tópicos de investigación como indexamiento en espacios métricos y no métricos, búsqueda por similitud en colecciones de objetos 3D, búsqueda de imágenes en la Web y teoría de indexamiento multimedia.



Grupo PRISMA: Benjamin Bustos, Violeta Chang, José Saavedra, Iván Sipirán y Juan Manuel Barrios.

Mis principales colaboradores en investigación son: Prof. Tomas Skopal, Charles University in Prague, República Checa; Dr. Tobias Schreck, Technische Universitaet Darmstadt, Alemania; Dr. Oscar Pedreira, Universidade da Coruña, España; Dra. Bárbara Poblete, Yahoo! Research Lab; Dr. Nelson Morales, DELPHOS Lab, AMTC, Universidad de Chile.

Alumnos de Posgrado

Actualmente dirijo cuatro estudiantes de Doctorado en Ciencias, mención Computación (Juan Manuel Barrios, José Saavedra, Iván Sipirán, y Violeta Chang, ésta última en conjunto con el profesor Gonzalo Navarro), y un alumno de Magister en Ciencias mención Computación (Víctor Sepúlveda).

Journals y Conferencias

En los últimos cinco años he publicado siete artículos de revista, 16 artículos en conferencias internacionales y dos capítulos de libro. Principalmente publico en las siguientes revistas y conferencias internacionales: ACM Computing Surveys; IEEE Transactions on Knowledge and Data Engineering; Multimedia Tools and Applications; Eurographics Workshop on 3D Object Retrieval (3DOR); International Conference on Similarity Search and Applications (SISAP).

Grupo de investigación

Soy Director del Grupo de Investigación PRISMA (Pattern Recognition, Similarity Search, and Indexing in Multimedia Archives), perteneciente al DCC de la Universidad de Chile. El objetivo principal del grupo es investigar nuevos algoritmos y técnicas para poder realizar búsquedas en grandes colecciones de datos multimedia en forma eficaz y eficiente.

En la actualidad, el grupo PRISMA trabaja en variados proyectos de investigación, que corresponden principalmente a las tesis de doctorado de los asistentes de investigación del grupo. Algunos de estos proyectos son: búsqueda en colecciones de modelos 3D; búsqueda con medidas de similitud no métricas; detección de copia de videos; búsqueda en imágenes basada en sketches; búsquedas por similitud usando índices comprimidos.

Desarrollo industrial y transferencia tecnológica

A través del Grupo de Investigación PRISMA, recientemente hemos realizado un exitoso proyecto de cooperación con la empresa chilena Orand, especializada en el desarrollo de software para proyectos de innovación. El proyecto consistió en el desarrollo de algoritmos para el reconocimiento del nombre y endoso en cheques manuscritos.

Esta tecnología se encuentra actualmente implementada en el "Chequemático", una máquina pagadora de cheques del Banco BCI. Actualmente se encuentran otros proyectos en carpeta para ser realizados junto a Orand.

Contacto

E-mail de contacto:
bebustos@dcc.uchile.cl.

Web del Grupo PRISMA:
<http://prisma.dcc.uchile.cl>.

Claudio Gutiérrez:

SEMÁNTICA, BASES DE DATOS, WEB

*Departamento de Ciencias de la Computación,
Universidad de Chile.*

Desde hace casi diez años, con diferentes colegas, hemos venido desarrollando en el Departamento de Ciencias de la Computación de la Universidad de Chile, los aspectos semánticos de manejo de datos en la Web.

Explicemos. Lo que hizo popular a la Web fue la aplicación de técnicas de recuperación de información, tradicionalmente un área completamente disjunta de las de bases de datos. La primera, anclada en técnicas estadísticas; la otra, en la lógica. Una tiene como objetivo recuperar la mayor cantidad (recall) de la mejor (según algún criterio) (precisión) información con poca estructura (lenguaje natural, documentos, etc.). La otra, responder lógicamente a consultas y razonar sobre la información estructurada. No es casualidad que ambas comunidades tengan poco en común.

El punto de partida fue la aplicación de técnicas clásicas de bases de datos (pensadas y motivadas por aplicaciones de negocios y empresariales) al ámbito de la Web. El gran inspirador de este enfoque fue Alberto



Carlos Hurtado, Alberto Mendelzon, asador, Claudio Gutiérrez y Gonzalo Navarro.

Mendelzon, quien era uno de principales teóricos de las bases de datos relacionales, un argentino muy latinoamericanista, que trabajaba en la Universidad de Toronto, en Canadá. Nuestro grupo tuvo la oportunidad de interactuar con él. Carlos Hurtado había sido su alumno en Toronto y por medio de él comenzamos a trabajar conjuntamente en estos temas.

Así comenzó a desarrollarse una masa crítica de investigadores y alumnos en torno a estos temas. El punto de partida fue estudiar RDF (Resource Description Framework), el lenguaje para describir recursos en la Web, como un modelo de datos, en la tradición de la disciplina de bases de datos. Partimos trabajando con Carlos Hurtado, con Alberto y luego con un conjunto amplio de colegas y estudiantes: Ernesto Krsulovic (estudiante de Magíster en el DCC de la Universidad de Chile, hoy consultor independiente), Renzo Angles (estudiante de Doctorado del DCC de la Universidad de Chile, hoy en la Universidad de Talca), Marcelo Arenas (Pontificia Universidad Católica, PUC), Jorge Pérez (Universidad de Talca, hoy terminando su doctorado en la PUC), Sergio Muñoz (Facultad de Ciencias de la Universidad de Chile), Alejandro Vaisman (Universidad de Buenos Aires), Andrea Rodríguez (Universidad de Concepción),

y muchos alumnos: Marcela Calderón, Cristián Vásquez, Álvaro Graves, Mauro San Martín, Daniel Hernández, etc. A nivel internacional nos acompañaron los profesores Leopoldo Bertossi (Toronto), Peter Wood (UK), Mariano Consens (Toronto), Axel Polleres (Irlanda), Enrico Franconi (Bolzano), Asunción Gómez-Pérez (Madrid), Manolis Koubourakis (Grecia), y varios otros. Y varios estudiantes que han venido del extranjero a visitar nuestro grupo y trabajar con él: Draltan Marín (el primero que especificó formalmente la semántica lógica de RDF), J. Hayes (que desarrolló el formalismo de grafos de RDF), Javier Fernández (que se ha dedicado a desarrollar la escalabilidad del formato RDF), etc.

El grupo desarrolló aspectos teóricos y prácticos de estos temas: Las especificaciones del Consorcio de la Web (W3C) en estas materias: RDF, RDFS, SPARQL; especificaciones para el Gobierno chileno (XML, metadatos, hoy DataGov). A nivel académico interactúa con grupos de Bases de Datos y de Web Semántica. Entre ellos están centros europeos, norteamericanos y latinoamericanos. Podemos señalar la Universidad de Buenos Aires y Bahía Blanca en Argentina, Universidad de la República en Uruguay, Universidad Católica de Arequipa en Perú, Universidad Central

de Venezuela, la PUC de Río de Janeiro. En Europa desarrollamos intercambio con DERI (Irlanda), UPM (España), Bolzano (Italia), TU Vienna (Austria), Oxford (UK), y en Estados Unidos el RPI. En la misma línea publica en conferencias de esa áreas: International Semantic Web Conference, Extended Semantic Web Conference, World Wide Web Conference, PODS, y diversos Workshops del área, y journals de bases de datos y Web Semántica, como TODS, TKDE, JWS, JCSS, etc.

En la actualidad el grupo está enfocado en el desarrollo de estos temas ligados a Linked Data, Open Data y movimientos que tienden a desarrollar los aspectos de razonamiento y escalabilidad en la Web. Entre las principales líneas de trabajo y actividad están:

- a) Desarrollo y estudio de estándares W3C: RDF, SPARQL.
- b) Participación formal e informal en grupos trabajo de W3C.
- c) Desarrollo y estudio de nuevos modelos de datos y lenguajes de consulta para la Web. Particularmente en torno a la especificación RDF: SPARQL, RDB2RDF, HDT.
- d) Desarrollo y estudio de aplicaciones de estas técnicas en Gobierno (DataGov) en la región y en Chile.
- e) Formación de comunidad, a través de seminarios, workshops y charlas o de investigación dirigidas a la comunidad local.
- f) Cursos y Extensión: A través de Educación Continua del DCC, en Cursos internacionales en Escuelas de Verano (Bolzano, UPM, Buenos Aires, Bahía Blanca, Montevideo, Arequipa, etc.).
- g) Contacto e intercambio con otras organizaciones y grupos de investigación en diferentes niveles (W3C, ONG locales, Gobierno, KHIPU, Datos-Chile, etc.)

Loreto Bravo:

LIMPIEZA Y CONSISTENCIA DE LOS DATOS

Departamento de Ingeniería Informática y Ciencias de la Computación, Universidad de Concepción.

Desde que comencé mi Doctorado me he centrado en temas de investigación relacionados con Datos Inconsistentes. Durante mis estudios en Canadá me centré, junto con mi supervisor, Leopoldo Bertossi, en el manejo de inconsistencias en bases de datos relacionales, en sistemas de integración de bases de datos y en sistemas P2P. En el contexto de bases de datos relacionales, nos concentramos en el problema de Consistent Query Answering y en la utilización de programas lógicos de reparación para computar las respuestas consistentes. Aplicando ideas de esta investigación, estudiamos además la semántica de sistemas de integración de datos y P2P en la presencia de restricciones de integridad.

Al terminar mi Doctorado realice un Posdoctorado en el grupo de Bases de Datos de la University of Edinburgh, UK. Ahí trabajé con Wenfei Fan, Floris Geerts y Shuai Ma en extensiones a restricciones de integridad, como dependencias funcionales y de inclusión, especialmente diseñadas para la limpieza de datos. Estudiamos los problemas de satisfacibilidad e implicancia para estas restricciones.

Durante mi Posdoc comencé también a trabajar, junto a Irini Fundulaki (ICS-Forth, Grecia) y James Cheney (University of Edinburgh, UK) en control de acceso para bases de datos XML. En particular, nos concentramos en la detección de inconsistencias de las políticas de control de acceso, es decir, en detectar si es posible conseguir por medio de una secuencia de operaciones permitidas una acción que



Loreto Bravo.

es prohibida. También hemos estudiado el problema de reparar las políticas en forma automática. Esta investigación ahora cuenta con el financiamiento de un proyecto Fondecyt de iniciación.

Ya instalada en Chile he comenzado a realizar investigación junto con Andrea Rodríguez (Universidad de Concepción) en la formalización y estudio de propiedades de restricciones de integridad para bases de datos espaciales. También he trabajado con Mónica Caniupán (Universidad del Bío-Bío), Carlos Hurtado (Universidad Adolfo Ibáñez) y Leopoldo Bertossi en consistencia de dimensiones de Data-Warehouses. Finalmente, también en conjunto con Leopoldo Bertossi, hemos continuando con la investigación de bases de datos P2P comenzada durante mi Doctorado y estamos preparando un artículo "Database Repairs and Consistent Query Answering" para la Synthesis Lectures on Data Management de Morgan & Claypool.

Los resultados de mi investigación han sido publicados en conferencias como VLDB, ICDE, EDBT, LPAR, DBPL, IJCAI y en las revistas Information Systems y Journal of Applied Logic de Elsevier.

Marcelo Arenas:

INTEROPERABILIDAD EN SISTEMAS DE MANEJO DE INFORMACIÓN

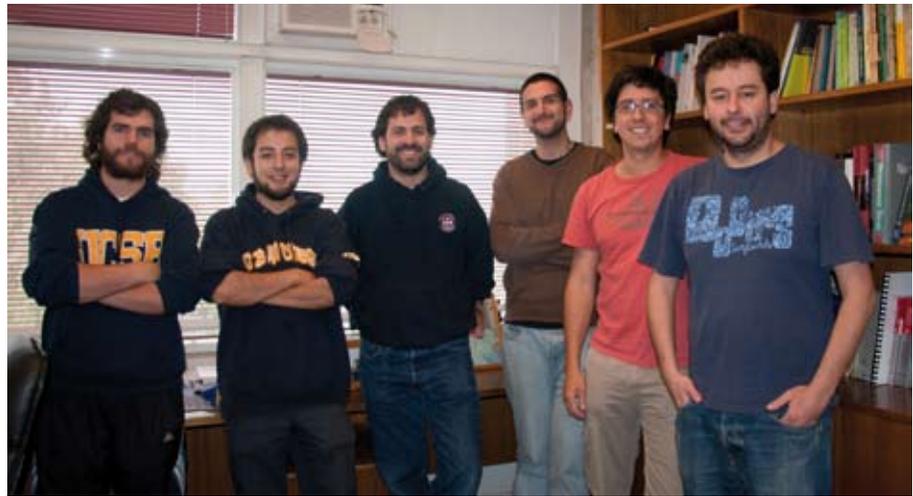
*Departamento de Ciencia de la Computación,
Pontificia Universidad Católica de Chile.*

El área de Bases de Datos, a pesar de ser un tema clásico en Ciencia de la Computación, cobra mucha relevancia hoy en día por los desafíos que imponen las nuevas tecnologías. Siguiendo esta premisa nuestro grupo investiga temas de manejo de información motivados por problemas de interoperabilidad entre aplicaciones muy relevantes por el uso de Internet y la Web.

Dos de nuestras principales áreas de investigación son el intercambio de información y la integración de información. El problema de intercambio de información surge cuando dos aplicaciones (bases de datos, páginas Web, servicios Web, etc.) que trabajan de manera independiente desean compartir información y a la vez mantener su independencia. Por su parte, en integración de datos el problema principal es proveer a un usuario (persona o máquina) de una vista unificada a fuentes de datos dispares e independientes. Ambos problemas están muy relacionados y varios de nuestros artículos han ayudado a formalizarlos y dar solución a algunos de los desafíos que ellos presentan.

Otra de nuestras áreas de investigación es el manejo de información en la Web semántica. Ésta es una iniciativa de la W3C para agregar información a la Web que tenga tanto sentido para las personas como para las máquinas. Nuestro grupo, colabora estrechamente con investigadores de la Universidad de Chile, ha estado en el centro de la definición de las tecnologías básicas de la Web semántica, en particular de los lenguajes de consulta para datos semánticos de la Web.

Una característica definitoria de nuestro grupo es la rigurosidad, tanto en la formalización de los problemas como en



Marcelo Arenas y su grupo PUC Chile.

la formulación de soluciones. Creemos firmemente que una base matemática sólida es esencial para dar soluciones que puedan ser comprobadamente mejores que las actuales y robustas de implementar. Es así como nuestra investigación tiene un fuerte componente teórico basado en herramientas como lógica computacional, en particular teoría de modelos finitos, complejidad computacional y complejidad descriptiva.

Parte de la calidad e impacto de nuestro trabajo puede ser medido por los premios académicos que estos han obtenido. Nuestro grupo ha obtenido cinco premios al mejor artículo ("Best Paper Award") en las más destacadas conferencias de teoría de bases de datos (PODS'03, PODS'05, ICDT'10) y Web semántica (ISWC'06, ESWC'07).

Colaboradores

Mantenemos una estrecha colaboración con investigadores de la Universidad de Chile, en particular con Pablo Barceló y Claudio Gutiérrez. Se destaca también nuestra colaboración con la industria internacional, en particular con Ron Fagin de IBM Almaden y Phil Bernstein de Microsoft Research. Parte de nuestros miembros han hecho pasantías y estadias cortas en estas empresas.

Adicionalmente colaboramos con investigadores de universidades internacionales entre los que podemos

destacar a Leonid Libkin, Juan Reutter, Wenfei Fan y Kousha Etessami, (University of Edinburgh); Juan Sequeda (University of Texas at Austin); Cristian Riveros (Oxford University); Axel Polleres (National University of Ireland); Leopoldo Bertossi (Carleton University); Mariano Consens (University of Toronto); Filip Murlak (University of Warsaw); Alan Nash (Aleph One LLC); Rajeev Alur (University of Pennsylvania); Neil Immerman (University of Massachusetts).

Alumnos vigentes

Jorge Pérez (Doctorado), Martín Ugarte (Doctorado), Carlos Buil-Aranda (Doctorado visitante, Universidad Politécnica de Madrid), Sebastián Conca (Magíster), Andrés Letelier (Magíster) y Alejandro Mallea (pregrado).

Conferencias internacionales

Publicamos en los últimos cinco años en las principales conferencias de bases de datos: ACM Symposium on Principles of Database Systems (PODS), International Conference on Database Theory (ICDT), e International Conference on Very Large Data Bases (VLDB). Publicamos también en las conferencias más importantes de Web semántica: International Semantic Web Conference (ISWC) y European Semantic Web Conference (ESWC). Parte

de nuestra investigación ha sido publicada en conferencias de lógica y autómatas como: Annual IEEE Symposium on Logic in Computer Science (LICS), y el International Colloquium on Automata, Languages and Programming (ICALP).

Revistas

En los últimos cinco años hemos publicado en las revistas: Journal of the ACM (JACM), SIAM Journal on Computing (SICOMP), ACM Transactions on Databases Systems (TODS), IEEE Transactions on Knowledge and Data Engineering (TKDE), SIGMOD Record, Annals of Pure and Applied Logic (APAL), Theory of Computing Systems (TOCS), Journal of Web Semantics (JWS), Logical Methods in Computer Science (LMCS), Journal of Computer and System Sciences (JCSS).

Número de artículos publicados en los últimos cinco años:

- Revistas: 14
- Conferencias internacionales: 11
- Libros: 1
- Capítulos de libros: 2
- Workshops internacionales: 3

Mónica Caniupán:

CONSISTENCIA DE DATOS SOBRE DIFERENTES MODELOS

Departamento de Ingeniería Civil Informática, Universidad del Bío-Bío.

Obtuve el grado de PhD in Computer Science en Carleton University (Ottawa, Canadá) el año 2007 bajo la supervisión del Dr. Leopoldo Bertossi. Mis intereses en investigación están centrados en: (a)

Teoría de Bases de Datos, (b) Integridad de Bases de Datos, (c) Calidad de Datos, (d) Representación del Conocimiento y (e) Programación Lógica.

Durante mi Doctorado me dediqué a estudiar y definir optimizaciones para programas en lógica de manera de ser utilizados en el cómputo de información consistente desde bases de datos inconsistentes (bases de datos que no satisfacen sus restricciones de integridad). La tesis se tituló "Optimizing and Implementing Repair Programs for Consistent Query Answering in Databases". Los resultados de esta investigación, fueron publicados parcialmente en: (a) In Current Trends in Database Technology, LNCS 3268 (2004), (b) Conferencia Internacional de la Sociedad Chilena de Computación" (2005), (c) "The Scalable Uncertainty Management Conference (SUM'07), LNCC 4772" (2007). Finalmente se publicó un artículo en Data and Knowledge Engineering Journal el año 2010 (69(6):545-572).

En 2007 me adjudiqué el proyecto Fondecyt de iniciación en investigación "Semantically Correct Answers to Queries in Inconsistent Multidimensional Databases". El objetivo de este proyecto fue definir una semántica de reparación para bases de datos multidimensionales que no satisfacen sus restricciones de integridad. Esto permitiría responder adecuadamente (consistentemente) a consultas de agregación. Demostramos en este trabajo que la teoría definida para bases de datos relacionales no puede ser aplicada a bases de datos multidimensionales; definimos una nueva semántica de reparación y una solución basada en programas en lógica para obtener las reparaciones minimales de dimensiones en Data Warehouses. En este proyecto colaboraron: Loreto Bravo (Universidad de Concepción), Carlos Hurtado (Universidad Adolfo Ibáñez) y Leopoldo Bertossi (Carleton University, Universidad de Concepción). El principal resultado de esta investigación fue enviado al "Data and Knowledge Engineering Journal" y en estos momentos se encuentra en proceso de revisión. También hemos publicado dos artículos en distintas versiones del



Mónica Caniupán.

"Alberto Mendelzon International Workshop on Foundations of Data Management" (2009,2010). Actualmente estoy trabajando junto con Alejandro Vaisman (Universidad de Buenos Aires) en la implementación de soluciones algorítmicas (no basadas en programación lógica) para obtener reparaciones de dimensiones en Data Warehouses. En esta investigación también participa la alumna Noemí Castillo del Magíster en Ciencias de la Computación de la Universidad del Bío-Bío, cuyo título de tesis es "Algoritmos para Computar Reparaciones de Dimensiones en Data Warehouses".

Además, he trabajado en manejo de inconsistencias en Bases de Datos Espaciales con Andrea Rodríguez (Universidad de Concepción) y Leopoldo Bertossi. Resultados parciales de esta investigación fueron publicados en "The 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS)" (2008). También enviamos un artículo a "Information System Journal" el cual se encuentra en proceso de revisión.

En los últimos cinco años he publicado cinco artículos en conferencias, un artículo en revista ISI y esperamos respuesta de dos artículos enviados a revistas ISI.



Grupo de investigación en Computación de Alto Rendimiento para la Web.

Mauricio Marín:

TEORÍA Y PRÁCTICA EN COMPUTACIÓN ESCALABLE PARA LA WEB

*Departamento de Ingeniería Informática,
Universidad de Santiago de Chile.*

El grupo de investigación en Computación de Alto Rendimiento para la Web se especializa en el estudio de problemas ubicados en la intersección entre recuperación de información, minería de datos, y procesamiento paralelo y distribuido de la información. Su objetivo principal es desarrollar estrategias que le permitan a las aplicaciones de la Web escalar a millones de usuarios de manera eficiente en términos de uso de recursos de hardware y software.

El financiamiento para investigadores y tesis de posgrado proviene de Yahoo! Research Latin America, el cual es un laboratorio de investigación alojado en la Fundación para la Transferencia Tecnológica (UNTEC) de la Universidad de Chile. A este laboratorio también concurren investigadores de varias universidades nacionales y uno de sus objetivos principales es convertirse en un referente en investigación aplicada trabajando en estrecha colaboración con los programas de posgrado nacionales. Otra línea de financiamiento proviene de Fondec en proyectos tales como el denominado "Observatorios Escalables de la Web

en Tiempo Real", en el cual participan académicos y estudiantes de la Universidad de Santiago, Universidad de Concepción, Universidad de Chile y Universidad Técnica Federico Santa María. También existen proyectos de inserción de capital humano avanzado de Corfo y Conicyt, los cuales posibilitan la inclusión de posdoctorandos en las líneas de investigación del grupo.

Algunos de los problemas de investigación estudiados tienen la siguiente motivación: se estima que actualmente los centros de datos contienen del orden de los 60 millones de computadores, los cuales consumen al menos el 2% de la energía a nivel mundial que se utiliza para generar electricidad. Por otra parte, la Web duplica su tamaño cada seis u ocho meses y aún faltan grandes sectores de la población mundial por incorporarse como usuarios de las diversas aplicaciones de la Web. Es, por tanto, relevante desarrollar estrategias que permitan a los centros de datos administrar a centenas de miles de usuarios concurrentes por segundo y a la vez sean eficientes en consumo de energía.

Típicamente los centros de datos operan sus computadores en régimen permanente a una utilización que está entre un 20% y un 40% de su capacidad total. La razón es que estos sistemas deben estar preparados para enfrentar subidas bruscas en el tráfico de peticiones de servicio de usuarios tales como consultas frente a eventos globales que capten el interés de cientos de miles de usuarios concurrentes por segundo. Una

línea de investigación desarrollada por el grupo tiene relación con el desarrollo de estrategias de procesamiento de consultas que sean capaces de reducir la cantidad de computadores desplegados en el centro de datos y hacerlos operar a una utilización mayor, pero incluir en ellos técnicas que les permitan absorber eficientemente subidas bruscas en el tráfico de consultas. Las técnicas desarrolladas tienen que ver con estrategias de caching e indexación distribuida, procesamiento paralelo de consultas tanto en sistemas de memoria distribuida como memoria compartida, y selección automática de nodos procesadores basada en aprendizaje de máquina.

El contacto con investigación aplicada real para sistemas Web de gran escala proviene por la vía de proyectos de investigación orientados al estudio de optimizaciones de productos de Yahoo! operando en producción. Actualmente se trabaja en dos proyectos relacionados con motores de búsqueda verticales. El primero, tiene relación con planeación de capacidad en el centro de datos, lo cual requiere el desarrollo de simuladores tanto a nivel macroscópico, es decir, simulación de *clusters* de nodos procesadores, como a nivel microscópico, esto es, simulación de procesadores multicore. Sobre estos simuladores que modelan el hardware, se desarrollan simuladores del software que componen los distintos servicios del motor de búsqueda vertical. Los desafíos en investigación están en la formulación de modelos pertinentes y su combinación con la aplicación de técnicas de optimización metaheurística orientadas a planificar el despliegue de servicios en los nodos procesadores del centro de datos. El segundo proyecto tiene relación con el empleo de técnicas de compresión de índices invertidos y multithreading, para hacer que los nodos procesadores que resuelven consultas enviadas al motor vertical, tengan capacidades de actualización en tiempo real de los documentos indexados en cada nodo. Los desafíos en investigación están en el desarrollo de técnicas de eficientes de indexación y gestión de threads para posibilitar la ejecución concurrente de transacciones de lectura y escritura sobre el índice comprimido.

Anualmente uno de los indicadores principales de desempeño del grupo de investigación tiene relación con la publicación de artículos en las conferencias más relevantes del área de recuperación de información y computación paralela y distribuida, como lo son las conferencias con acrónimos SIGIR, WWW, CIKM, ECIR, SPIRE, HPDC, ICPP, IPDPS y Euro-Par. También es relevante generar patentes en los Estados Unidos de América.



Pablo Barceló.

Pablo Barceló:

MODELOS EMERGENTES DE DATOS

Departamento de Ciencias de la Computación, Universidad de Chile.

Desde los años de mi Doctorado, realizado entre 2002 y 2006 en el Departamento de Ciencia de la Computación de la Universidad de Toronto, Canadá, vengo realizando investigación en modelos emergentes de representación y consulta de datos. Esto se refiere principalmente a dos cosas:

(1) El estudio de nuevos formatos para el manejo de información, que van más allá del tradicional modelo relacional, y que han sido impuestos por la aparición en los últimos 15 años de aplicaciones centradas en datos tan importantes como la Web, las bases de datos científicas, las redes sociales, entre varias otras. Estos nuevos formatos de datos destacan por permitir mayor flexibilidad de representación que el modelo relacional, manteniéndose al mismo tiempo la posibilidad de entregar cierta estructura a partir de elementos semánticos y jerarquías. Por esta razón se han llamado “semiestructurados” a este tipo de datos.

En particular, mi investigación se ha centrado en torno a dos modelos de datos semiestructurados: (a) XML (Extensible Markup Language), que es un metalenguaje que permite describir información a alto nivel, y que se ha convertido en el estándar para integrar e intercambiar información en la Web; y (b) las bases de datos de grafos, que es un modelo abstracto que se utiliza para describir aplicaciones centradas en los

datos en las que la que la topología de estos es tan importante como los datos mismos (por ejemplo, redes sociales, bases de datos científicas, Web semántica, etc.) En ambos modelos de datos mi investigación se ha centrado en torno a el diseño y análisis de lenguajes de consulta (por ejemplo, entender la expresividad y complejidad de evaluación de estos) y la representación y estudio de la información incompleta e incierta (que aparece ubicuamente en escenarios distribuidos como la Web, donde la información está fragmentada y podría presentar altos grados de incertidumbre).

(2) El estudio de problemas dinámicos asociados a los datos, que aparecen en espacios en donde la información fluye constantemente como la Web, y que no corresponden a la línea más tradicional de estudio en bases de datos donde estos son considerados estáticos. Me he enfocado, en particular, en estudiar problemas dinámicos asociados a la integración y el intercambio de la información. Mi investigación se ha centrado en entender la complejidad computacional de los distintos problemas relacionados con estos dos temas, así como en la potencial aplicación de los lenguajes de consulta tradicionales –por ejemplo, SQL– en este escenario más complejo.

Aunque mi formación de pregrado no es en Ciencia de la Computación –soy Ingeniero Electricista de la Universidad Católica de Chile– siempre me han atraído los temas de la Computación, en particular aquellos que tienen que ver con la teoría y los algoritmos. En particular, los temas de bases de datos concitaron desde un primer momento mi atención porque combinan, de forma bastante equilibrada, dos de mis intereses:

- (1) La posibilidad de realizar modelos abstractos de los datos, que no dependieran de una aplicación en particular, sino que más bien aglutinaran las características esenciales que definen a una familia de aplicaciones. Estos modelos abstractos se prestan naturalmente al análisis lógico/matemático de alto nivel, combinando de forma muy interesante herramientas que van desde teoría de autómatas, pasando por expresividad de lenguajes lógicos, hasta llegar a la teoría de complejidad. Muchas veces el trabajo matemático que se hace en bases de datos es de alta dificultad, no teniéndole nada que envidiar al análisis que se hace en otras ramas más teóricas de la computación.
- (2) La posibilidad de que dicho estudio teórico sea de impacto para la comunidad más aplicada. Es decir, las bases de datos son un interesante espacio de problemas para el teórico, pero a la vez proveen el espacio para descubrir, mediante dicho estudio, propiedades fundamentales de los modelos de datos que pueden ayudar a la comunidad más aplicada a desarrollar aplicaciones más robustas y eficientes.

Colaboradores

Como es usual en Ciencia de la Computación, nuestro trabajo se ha desarrollado en cercana colaboración con investigadores a lo largo del mundo. En Chile mantengo cercano contacto con Marcelo Arenas (Pontificia Universidad Católica). Mis más cercanos grupos de investigación en la actualidad son

la Universidad de Edinburgo, donde trabajo con el profesor Leonid Libkin y su alumno de Doctorado Juan Reutter, y la Universidad de Ottawa, donde me encuentro trabajando con el Profesor Iluju Kiringa. Además he publicado libros y artículos con Leopoldo Bertossi (Carleton University) Ron Fagin (IBM, Almaden), Rajeev Alur (University of Pennsylvania), Kousha Etessami (University of Edinburgh), Neil Immerman (University Massachusetts), Loreto Bravo (Universidad de Concepción), Cristina Sirangelo (INRIA), Antonella Poggi (University of Rome), Filip Murlak (University of Warsaw), Carlos Hurtado (Universidad Adolfo Ibañez) y Peter Wood (University of London).

Publicaciones

Parte de nuestro trabajo calza perfectamente con el perfil de estudio tradicional de fundamentos de bases de datos – como, por ejemplo, lo que tiene que ver con diseño de lenguajes de consulta, integración de información o información incompleta – y, por tanto, usualmente lo publicamos en las conferencias estándares del área como son ACM Principles of Database Systems (PODS) e International Conference on Database Theory (ICDT). Otra parte de nuestro trabajo es más teórica, y tiene que ver con las propiedades lógicas esenciales de las estructuras que subyacen a nuestros modelos de datos. Como tal, usualmente publicamos ese tipo de trabajos en las conferencias teóricas del área de lógica y autómatas como IEEE Logic in Computer Science (LICS), International Colloquium on Automata, Languages and Programming (ICALP) y Computer Science Logic (CSL).

Además, nuestros artículos han sido publicados en las siguientes revistas internacionales: Journal of the ACM (JACM), Sigmod Record, Annals of Pure and Applied Logic (APAL) y Theory of Computing Systems (ToCS). Recientemente hemos también publicado el libro “Relational and XML Data Exchange”, junto a Marcelo Arenas, Leonid Libkin y Filip Murlak, en las Synthesis Lectures on Data Management de la editorial Morgan & Claypool.

Sistemas Colaborativos + Interacción Humano Computador

CARL: COLLABORATIVE APPLICATIONS RESEARCH LABORATORY

Departamento de Ciencias de la Computación, Universidad de Chile.

En este laboratorio de investigación participan tres académicos: Nelson A. Baloian, Sergio F. Ochoa, José A. Pino y diversos tesis de Doctorado y Magíster del Departamento de Ciencias de la Computación de la Universidad de Chile.

Tal como lo indica el nombre del laboratorio, el trabajo de investigación aquí realizado está enfocado en los Sistemas Colaborativos (SC). Este tipo de sistemas involucra componentes de software y hardware que apoyan el trabajo de equipos de personas que cooperan para alcanzar un objetivo común. Algunos ejemplos de sistemas colaborativos bien conocidos son las agendas compartidas, los foros de discusión y las redes sociales. Sin embargo, hay un sinnúmero de aplicaciones para áreas específicas, donde este tipo de sistemas tiene un rol protagónico, por ejemplo en la toma de decisiones, en escenarios educativos, en trabajo móvil débilmente acoplado y en el procesamiento de flujos de trabajo (workflows), entre otros.

Este tipo de sistemas contrasta con el enfoque tradicional de provisión de herramientas computacionales para el trabajo individual.

También contrasta con enfoques en los cuales la máquina es quien resuelve problemas; aquí el énfasis está en potenciar a las personas para que ellas mismas resuelvan los problemas.

Dentro de los SC, los investigadores de CARL se han concentrado especialmente en temas de colaboración móvil de apoyo a diversas áreas de trabajo, como por ejemplo: la respuesta a emergencias, las inspecciones de obras civiles, el trabajo en hospitales, el aprendizaje colaborativo y la evaluación del apoyo que brindan al usuario estos sistemas. Muchos de los proyectos en desarrollo abarcan más de uno de estos temas, y son ejecutados en conjunto con investigadores extranjeros. Particularmente CARL mantiene una estrecha colaboración con laboratorios similares de la Universidad de Lisboa (Portugal), Universidad Federal de Río de Janeiro (Brasil), CICESE (México) y Universidad de Duisburg-Essen (Alemania). El financiamiento externo de apoyo al trabajo científico en estas áreas proviene de diversas fuentes, como por ejemplo: Fondecyt, LACCIR, CYTED, DAAD y proyectos europeos.

Entre los logros obtenidos por investigadores de CARL se cuentan framework, modelos y técnicas de diseño y evaluación de SC, al igual que infraestructuras de apoyo al trabajo colaborativo móvil. También se han desarrollado modelos y software de apoyo a Bomberos en su respuesta a emergencias, a personal médico que



Nelson Baloian, José A. Pino y Sergio F. Ochoa.

trabaja en hospitales y a inspectores de obra en proyectos de construcción. En el área de aprendizaje colaborativo, se ha trabajado en el apoyo computacional dentro de la sala de clase, en la integración del aprendizaje dentro y fuera de la sala de clase, y en la inferencia de metadatos de objetos de aprendizaje. Anteriormente, se desarrolló un método para versionar objetos construidos colaborativamente (citado por dos patentes).

El grupo publica en revistas tales como: *Advanced Engineering Informatics*, *Expert Systems with Applications*, *Knowledge and Information Systems*, *Educational Technology and Society*, *ACM Computing Surveys*, *Personal and Ubiquitous Computing*, *Group Support Systems* y *Group Decision and Negotiation*. También publica en revistas que se enfocan en los aspectos estructurales de los sistemas colaborativos, como por ejemplo el *Journal of Network and Computer Applications*, and *Journal of Systems and Software*. En relación a conferencias internacionales, los miembros del laboratorio participan en eventos tales como CRIWG, CSCL, CSCWD, HICSS.

CENTRO DE COMPUTACIÓN Y COMUNICACIÓN PARA LA CONSTRUCCIÓN DEL CONOCIMIENTO, C5

Departamento de Ciencias de la Computación, Universidad de Chile.

El C5 (www.c5.cl) es un centro de investigación y desarrollo, cuyo objetivo es la creación de herramientas tecnológicas de punta para asistir y mejorar el aprendizaje y desarrollo intelectual de niños y jóvenes. En el Centro también, se evalúan y llevan a cabo tecnologías interactivas para la inclusión de personas con discapacidad visual. Está conformado por un equipo multidisciplinario de profesionales de las áreas de computación, educación, diseño gráfico y sociología, entre otras.

Para conocer más de este Centro, consultar sección "Grupos de Investigación" (página 82).

Computación Gráfica + Computación Científica

María Cecilia Rivara:

COMPUTACIÓN CIENTÍFICA, MODELACIÓN GEOMÉTRICA Y VISUALIZACIÓN: UN CONTEXTO INTERDISCIPLINARIO

Departamento de Ciencias de la Computación, Universidad de Chile.

Computación Gráfica es un tema complejo y dual en distintos sentidos. Sus comienzos en los años '60, están fuertemente relacionados con las aplicaciones de ingeniería. Fueron las industrias aeroespacial y de automóviles en Estados Unidos quienes financiaron tanto el desarrollo de computadores CRT, el desarrollo de software visual interactivo CAD (diseño asistido por computador) y el desarrollo de software de elementos finitos para análisis en ingeniería de problemas modelados por ecuaciones diferenciales parciales. El trabajo pionero de Iván Sutherland con su tesis de doctorado en 1963 sienta las bases de lo que sería la computación gráfica actual. Sin embargo, sólo fue en la década de los '80, con la masificación de los computadores personales *raster* que la

computación gráfica se consolida como un campo de conocimiento en sí mismo, con intensiva investigación interdisciplinaria, y aplicaciones masivas.

Fueron también los ingenieros los primeros en crear y estudiar métodos computacionales, estructuras de datos y algoritmos para construir discretizaciones (triangulaciones en 2D y 3D, mallas de cuadriláteros y hexaedros) como herramientas necesarias para el análisis de problemas físicos complejos mediante métodos de elementos finitos. Es sólo posteriormente que los investigadores de matemáticas estudian los métodos de elementos finitos, y los investigadores de Ciencias de la Computación y de Computación Gráfica estudian los algoritmos para construir triangulaciones y desarrollan técnicas específicas para los requerimientos del área.

En aplicaciones de ciencias, ingeniería y medicina, se define visualización científica al área interdisciplinaria del conocimiento cuyo objetivo es crear herramientas computacionales visuales para contribuir a la comprensión de enormes y complejos conjuntos de datos (por ejemplo, astronómicos, geofísicos y geológicos), así como a la comprensión y validación de los resultados de modelos computacionales complejos, requeridos en computación



Ma. Cecilia Rivara, Pedro Rodríguez, Francisca Gallardo y Víctor Toledo.

científica. Computación Científica se define actualmente como el área amplia que abarca el estudio y uso de los métodos numéricos útiles en la práctica, así como el desarrollo de software para simular, analizar y estudiar fenómenos aplicados complejos.

Investigación interdisciplinaria

Me formé como Ingeniero Matemático con especialización en análisis numérico en la Facultad de Ciencias Físicas y Matemáticas (FCFM) de la Universidad de Chile, realicé una tesis de Doctorado en desarrollo de software adaptivo para métodos de elementos finitos en la Katholieke Universiteit Leuven en Bélgica. Como parte de esta tesis propuse los primeros algoritmos de refinamiento de triangulaciones usados en aplicaciones de ingeniería. En 1989 me integré al DCC de la Universidad de Chile y me hice cargo de los cursos de computación gráfica y de los cursos de mallas geométricas. Actualmente trabajamos en el estudio de las propiedades teóricas de algoritmos de triangulaciones, en el desarrollo de algoritmos mejorados en dos y tres dimensiones, en la paralelización de estos algoritmos, en aplicaciones de computación científica y de visualización científica.

En los últimos cinco años he publicado siete artículos en las revistas: *Computer-Aided Design*, *Applied Numerical Mathematics*, *Advances in Engineering Software*, *Engineering with Computers*, *Theoretical Computer Science*, *Journal of Computational and Applied Mathematics*, y presentado trabajos en las conferencias *International Meshing Roundtable*, *European Conference on Numerical Mathematics and Advanced Applications*, *Geometric Modeling and Processing*, en el *Chilean Workshop on Numerical Analysis of Partial Differential Equations*, *ENIEF*, *GRAPP*. He colaborado con los profesores Claudio Gutiérrez del DCC de la Universidad de Chile, Bruce Simpson de la Universidad de Waterloo, Nikos Chrisochoides del *College of William and Mary*, Rafael Montenegro de la Universidad de Las Palmas de la Gran Canaria e investigadores de la Universidad de Girona. Participo como investigadora

asociada en el proyecto “Modelos Numéricos Predictores para Gestión Medioambiental” del Ministerio de Ciencia e Innovación de España liderado por Rafael Montenegro.

En los últimos cinco años han participado en estos temas los alumnos de posgrado Pedro Rodríguez, Marité Guerrieri (Universidad de Girona), David Azócar (Magister Ingeniería Mecánica FCFM), Oliver Vilca, Víctor Toledo, Carlo Calderón, Eduardo Graells, Ana Cáceres y Carlos Lillo. Han participado también los alumnos de pregrado Gastón Jorquera, Álvaro Faúndez, Carlo Calderón, Pedro Valenzuela, Bernd Bierdermann, Joon Kim y Tomás Henríquez.

Fernando Rannou:

HIGH PERFORMANCE COMPUTING IN MEDICINE AND BIOLOGY GROUP

Departamento de Ingeniería Informática, Universidad de Santiago de Chile.

Desde hace algunos años hemos ido formando un grupo pequeño, pero muy sólido, de personas interesadas en usar tecnologías computacionales para resolver diversos problemas en el área de la medicina, biología y otras afines.

Nos hacemos llamar el High Performance Computing in Medicine and Biology Group, y otras yerbas, para significar nuestro

interés en el uso aplicado de técnicas de computación paralela y distribuidas en salud y, principalmente, ciencias de la vida. A veces nos enfocamos más en la aplicación y otras veces más en la tecnología computacional, todo dependiendo de cómo se va dando la investigación, pero nunca nos alejamos demasiado de ambos ámbitos.

Por ejemplo, durante el proyecto Fondecyt anterior estudiamos métodos de reconstrucción tomográficos de imágenes a partir de datos de escáneres para animales pequeños. Se estudiaron algoritmos estadísticos/iterativos que prometen producir mejores imágenes que los algoritmos basados en imágenes transformadas. Usamos una herramienta de simulación Monte Carlo llamada GATE para modelar y simular experimentos realísticos de uno de los scanners Positron Emisión Tomography (PET). Así, podemos estimar la respuesta del sistema (scanner) y usar dicho modelo en la reconstrucción iterativa. En este trabajo cooperamos con el Crump Institute for Molecular Imaging (UCLA), quienes realmente construyen los scanners prototipos para ratones. También, estamos incursionando en diferentes técnicas de aceleración de estos algoritmos, como por ejemplo, algoritmos paralelos en procesadores convencionales, algoritmos paralelos en GPU, y algoritmos *cache aware*.

Actualmente estamos un poco más interesados en los aspectos computacionales de la herramienta GATE. En el proyecto actual



Fabrizio Barisione, Felipe Torres, Nicolás Vega, José Luis Allende, Pablo Torres y Fernando Rannou.

hemos propuesto paralelizar completamente dicha herramienta, pues las simulaciones pueden tardarse semanas o incluso meses. Sin embargo, la paralelización de una simulación GATE no es tan obvia, pues los experimentos que simula (PET) restringen las posibilidades y el rendimiento que se puede teóricamente alcanzar. En este sentido es de vital importancia conocer cómo funciona PET; no basta simplemente con dividir el trabajo en los nodos.

El OpenGATE Collaboration es un grupo de investigadores que se encarga de mantener esta herramienta. Nosotros, como Universidad, formamos parte de dicha colaboración y también del GATE Steering Committee. El paper oficial de GATE fue premiado como la publicación más referenciada en el período 2004-2009 en la revista *Physics in Medicine and Biology*, reflejando lo importante que es esta herramienta en la comunidad de medicina nuclear.

Participamos en las conferencias IEEE Nuclear Science Symposium and Medical Imaging Conference y World Molecular Imaging Congress.

Publicamos en las revistas IEEE Transactions on Nuclear Science, *Physics in Medicine and Biology* y *Medical Physics*.

Como se dijo anteriormente, el grupo de investigación es pequeño. Actualmente está compuesto por un alumno de Magister y cuatro alumnos de Ingeniería Civil en Informática. En la foto aparecen de izquierda a derecha, Fabrizio Barisione (algoritmos iterativos secuenciales *cache aware*), Felipe Torres (paralelización de algoritmos iterativos), Nicolás Vega (ex alumno, colaborador externo), José Luis Allende (algoritmos iterativos en GPU), Pablo Torres (GATE multihebras), y Fernando Rannou (profesor). Como parte de su formación, todos los alumnos deben preparar y presentar al menos un paper en conferencias internacionales, nacionales o incluso revistas indexadas. Esta práctica nos ha permitido mantener una humilde tasa de una publicación journal por año y una en conferencia.

En el futuro esperamos extender el grupo para incluir alumnos de Doctorado en Ciencias de la Ingeniería mención Informática y también un posdoctorado. Si alguien está interesado en explorar las posibilidades de trabajar con nosotros, escriban a: fernando.rannou@usach.cl

Luis Salinas:

APLICACIONES DE COMPUTACIÓN EN CIENCIAS E INGENIERÍA

Departamento de Informática, Universidad Técnica Federico Santa María.

Director Centro de Innovación Tecnológica en Computación de Alto Desempeño UTFSM (CTI-HPC UTFSM).

Subdirector Centro Científico Tecnológico de Valparaíso.

Mi trabajo de investigación se centra en las aplicaciones de la computación en ciencias e ingeniería. Este es un campo eminentemente interdisciplinario, donde a menudo intervienen aspectos avanzados de computación, matemática, física, diversas ciencias de la ingeniería (dependiendo del

problema concreto estudiado) y, a veces, de disciplinas tan disímiles como la medicina (procesamiento de imágenes médicas) o de las finanzas. Los problemas relacionados con la computación, que surgen hoy en estas áreas están lejos de ser triviales y ciertamente no se reducen a una "simple" programación. Normalmente es necesario partir de un adecuado modelado matemático-físico del problema, que tenga en cuenta las particularidades del problema estudiado, hasta llegar a elección de las plataformas de software y hardware más apropiados, frecuentemente con fines de simulación o experimentación computacional.

Muy brevemente mis temas de investigación preferidos son la Computación Científica, la Computación de Alto Desempeño, Grid Computing, Finanzas Computacionales, Métodos Computacionales en Ingeniería, Ciencia y Matemática. Estos temas caen dentro del área de los Métodos Cuantitativos del Departamento de Informática de la Universidad Técnica Federico Santa María (UTFSM), que integro junto a los colegas Héctor Allende, Carlos Castro y María Cristina Riff.

Mis colaboradores más cercanos son Stephan Ruscheweyh (U. Würzburg, Würzburg, Alemania), Gonzalo Hernández (Centro de Modelamiento Matemático



Luis Salinas junto a su Grupo de Investigación.

(CMM), Universidad de Chile, Santiago, Chile), Alejandro Cañete (Innovative Financial Technology, New York-Santiago), Lorena Barba (Boston University, Boston, Estados Unidos), Óscar Orellana y Javier Cañas (UTFSM, Valparaíso, Chile). Ellos constituyen el núcleo académico del Centro de Innovación Tecnológica en Computación de Alto Desempeño (CTI-HPC) de la UTFSM, del cual soy su Director desde su fundación a fines de 2008. La UTFSM apoya el CTI-HPC en presupuesto e infraestructura; su misión es acercar la universidad a la industria y la empresa, y se espera que se autofinancie hacia el año 2015. El CTI-HPC cuenta con un staff estable de unos diez investigadores, usualmente estudiantes de posgrado de la UTFSM. El CTI-HPC es uno de los tres componentes del Centro Científico Tecnológico de Valparaíso (CCTVal), que es uno de los trece Centros Científicos y Tecnológicos de Excelencia patrocinados por el Programa de Financiamiento Basal de CONICYT. Los otros dos componentes del CCTVal son: Física (liderada por Iván Schmidt) y Electrónica de Potencia (liderada por José Rodríguez).

En el área de grid computing colaboramos estrechamente con el grupo de investigación avanzada en Física de Altas Energías de la UTFSM (I. Schmidt, W. Brooks, S. Kopeliovich, Y. Ivanov). Nuestro *cluster* computacional está integrado en la grid mundial del experimento ATLAS del Large Hadron Collider del CERN, Ginebra, Suiza. Por medio de la grid de ATLAS participamos activamente -en colaboración con los colegas de Física- en los trascendentales experimentos que actualmente se desarrollan en el LHC. Nuestros estudiantes de posgrado realizan a menudo pasantías de investigación en el CERN.

En los últimos diez años el grupo de investigadores del CTI-HPC ha publicado más de 100 trabajos en conferencias y journals internacionales, de primer nivel.

Entre las conferencias están: Symposia in HPC, International Conference on Computing in High Energy and Nuclear Physics, EELA Conferences, SCAT Conferences, Latin American Conference on High Performance Computing, Congresos SCCC, Grupo de Estudios en Ingeniería Clínica (GEIC), etc. Entre los journals donde el grupo publica, están: Applied Intelligence, Lecture Notes in Computer Science, Lecture Series on Computer and Computational Sciences, Journal of Mathematical Analysis and Applications, Physica A, Mathematische Zeitschrift, Israel Journal of Mathematics, AMS Contemporary Mathematics, Proceedings of the American Mathematical Society, Glasgow Mathematical Journal, American Institute of Physics Conference Proceedings, Numerical Analysis and Applied Mathematics, etc.

En la actualidad mis alumnos de Doctorado en Ingeniería Informática son Raquel Pezoa, Paola Arce, César Fernández, Roberto León; y mis alumnos de Magíster en Informática, Roberto Bonvallet y Rafael Plaza.

Mis proyectos de investigación más recientes incluyen: Computational and Geometric Function Theory 2010-2012 (FONDECYT 1100805); Computational and Geometric Function Theory 2007-2009 (FONDECYT 1070269); Computational and Geometric Function Theory (con S. Ruscheweyh, U. Würzburg, Alemania; FONDECYT 7070131, 7080064); High Performance Cluster Computing in Science and Engineering (DGIP-UTFSM 240954); Institute for advanced studies in science and engineering (Proyecto Anillo, CONICYT, con Will Brooks, UTFSM); Centro Científico Tecnológico de Valparaíso (Programa de Financiamiento Basal de CONICYT, proyecto FB 0821; con I. Schmidt y J. Rodríguez); Center for Technological Innovation: High Performance Computing in Science and Engineering, UTFSM (con Óscar Orellana y Javier Cañas); EPIKH: Exchange Programme to advance e-Infrastructure Know-How (Marie

Curie Actions-International Research Staff Exchange Scheme, European Community; con Roberto Barbera et al., U. Catania, Italia); EELA-2: E-Science grid facility for Europe and Latin America (ALFA Projects, European Community; con Bernard Marechal et al., CIEMAT, España); EELA: E-Infrastructure shared between Europe and Latin America (ALFA Projects, European Community; con Bernard Marechal et al., CIEMAT, España); SCAT: Scientific Computing and Advanced Training (ALFA Projects, European Community; con Lorena Barba, U. Boston).

Eventos organizados por el grupo en la UTFSM, Valparaíso, son: GISELA/EPIKH School for Grid Site Administrators (22-26 de Noviembre de 2010); CHAIN/GISELA/EPIKH School for Application Porting (29 de noviembre al 9 de diciembre de 2010); EPIKH Workshop (10 de diciembre 2010); Scientific Computing in the Americas: the challenge of massive parallelism, 3-14 January 2011, Valparaíso, Chile. Pan-American Advanced Studies Institute (NSF).

Nancy Hitschfeld:

MODELAMIENTO GEOMÉTRICO PARA APLICACIONES EN CIENCIA E INGENIERÍA

Departamento de Ciencias de la Computación, Universidad de Chile.

La investigación que realizo se enmarca en el área de la computación visual y gráfica, y consiste en el diseño e implementación de algoritmos para el modelamiento geométrico de objetos y análisis de imágenes en el contexto de aplicaciones en ciencia e ingeniería. En el desarrollo de las distintas herramientas han participado tanto alumnos de pregrado como de posgrado, la mayoría

del Departamento de Ciencias de la Computación (DCC) de la Universidad de Chile.

Dependiendo del área de aplicación de las distintas herramientas, he colaborado con profesores/investigadores de distintas áreas de la ciencia e ingeniería.

En particular, desde 2005 he estado realizando investigación en los siguientes temas:

- Desarrollo de un generador de mallas de superficie, de propósito general, pero también adaptado para el modelamiento de deformaciones de árboles y modelamiento de rostros. En el desarrollo del software trabajaron los alumnos Francisco Medina, Nicolás Silva, Cristina Melo y Renato Valenzuela (DCC Universidad de Chile), y en la paralelización de algunos algoritmos sobre GPU's, Cristóbal Navarro (Instituto de Informática, Universidad Austral). Para el modelamiento de deformaciones de árboles he trabajado en conjunto con el profesor Jaime San Martín (Centro de Modelamiento Matemático (CMM); Departamento de Ingeniería Matemática, Universidad de Chile), el investigador Fernando Padilla (CMM), y la profesora Eliana Scheihing (Instituto de Informática, Universidad Austral) y para el modelamiento de rostros con los investigadores Mauricio Cerda y Bernard Gireau (Inria-Loria, Francia), Lucas Terissi y Juan Carlos Gómez (Universidad de Rosario, Argentina). Este trabajo ha sido parcialmente financiado por un proyecto Fondecyt y un proyecto Stic-amSud (BAVI).
- Modelamiento de órganos humanos usando mallas mixtas y de hexaedros en conjunto con los profesores Claudio Lobos (Departamento de Informática, Universidad Técnica Federico Santa María) y Yohan Payan (TIMC-IMAG, Universidad Joseph Fourier), y el investigador Marek Bucki (TIMC-IMAG, Universidad Joseph Fourier). Este trabajo ha sido financiado parcialmente por



Pablo Aguilar, Diego Díaz, Nancy Hitschfeld y Jorge Jara.

- un proyecto Stic-amSud (PLOMO) y un proyecto ECOS-CONICYT.
- Desarrollo de algoritmos para aplicaciones en astronomía, en particular para la detección automática de *clusters* de galaxias y reconstrucción de imágenes, con los alumnos de posgrado Daniel Pizarro y Guillermo Cabrera (DCC Universidad de Chile), y los profesores Luis Campusano y Simon Casassus (Departamento de Astronomía, Universidad de Chile).
- Diseño de una arquitectura para familias de productos de generadores de mallas, con los profesores Cecilia Bastarrica (DCC Universidad de Chile) y Pedro Rossel (Departamento de Informática, UCSC); los alumnos Carlos Lillo, Felipe Contreras (DCC), y Jocelyn Simmonds (DCC Universidad de Chile, alumna de doctorado en la Universidad de Toronto).
- Análisis de imágenes, en particular aplicado a la detección de anillos de árboles, en colaboración con los alumnos Mauricio Cerda y Pablo Aguilar (DCC Universidad de Chile), el investigador Fernando Padilla (CMM), el profesor Domingo Mery (Departamento Ciencias de la Computación, Pontificia Universidad Católica).
- Desarrollo de algoritmos geométricos para aplicaciones en biología celular. En colaboración con el profesor Steffen Haertel, (SCIEN-LAB, Instituto de Ciencias Biomédicas, Universidad de Chile) y los alumnos de doctorado Jorge Jara y Pablo Aguilar (DCC Universidad de Chile).
- Desarrollo de un generador de mallas mixtas en tres dimensiones para la simulación de problemas basados en el método numérico de volúmenes finitos.

Considerando desde el año 2005, los resultados obtenidos han sido publicados (o in press.) en siete artículos de revistas internacionales: en *Engineering with Computers* los trabajos sobre generación de mallas; en *Advances in Engineering Software* y *Journal of Software Engineering and Knowledge Engineering*, los trabajos relacionados a ingeniería de software; en *Astrophysical Journal* los trabajos relativos con astronomía, y en *Insight* los relacionados con imágenes. Entre las conferencias donde se han publicado resultados están: *International Meshing Roundtable (IMR)*, *International Symposium on Voronoi Diagrams in Science and Engineering (ISVD)*, *Advances Image Video and Technology (PSIVT)*, e *International Conference on Software Reuse, (ICRS)*.

Sistemas Inteligentes

Carlos Castro, Eric Monfroy, Broderick Crawford:

SATISFACCIÓN RESTRINGIDA Y OPTIMIZACIÓN DISCRETA

*Departamento de Informática, Universidad
Técnica Federico Santa María.*

*Escuela de Ingeniería Informática, Pontificia
Universidad Católica de Valparaíso.*

En general, nos interesa la resolución de Problemas de Satisfacción de Restricciones y la Optimización Discreta utilizando técnicas completas, incompletas e híbridas que integren las dos primeras.

El estudio de las técnicas completas, incompletas e híbridas ha utilizado diversos problemas de planificación, en particular, hemos trabajado el Problema de Balanceo de Mallas Curriculares Académicas. Este problema propuesto originalmente por Carlos Castro, fue incluido en la librería de Problemas de Satisfacción de Restricciones www.csplib.org y se ha convertido en un problema de gran interés para la comunidad de Programación con Restricciones siendo actualmente objeto de diversos estudios y extensiones.

En el contexto de la integración de técnicas, trabajamos en el desarrollo de algoritmos que se adapten para determinar las mejores heurísticas a aplicar para un problema dado. La noción de Autonomous Search ha sido estudiada y los avances recientes en el tema serán publicados en 2011 por Springer Verlag en un libro donde Eric Monfroy es co-editor junto con Youssef Hamadi (Microsoft Research Cambridge, UK) y Frédéric Saubion (Universidad de Angers, Francia).

Por otro lado, trabajamos en la aplicación del enfoque basado en restricciones para la composición de servicios Web. Estamos interesados específicamente en



Carlos Castro.

los problemas de composición horizontal, donde la composición es fijada de manera abstracta. El problema consiste en determinar nuevas instancias de los servicios Web que permiten, por ejemplo, satisfacer la solicitud de un cliente. Este tipo de problema puede ser visto como un problema de configuración que puede ser tratado gracias a un razonamiento basado en restricciones. Los primeros resultados de este trabajo fueron presentados en la International Conference on Cooperative Information Systems en un trabajo conjunto con Olivier Perrin (Universidad de Nancy I, Francia), Christophe Ringeissen (INRIA, Francia) y Laurent Vigneron (Universidad de Nancy II, Francia).

Carlos Hernández:

BÚSQUEDA HEURÍSTICA EN TIEMPO REAL

*Departamento de Ingeniería Informática,
Universidad Católica de la Santísima
Concepción.*

Trabajo en el Departamento de Ingeniería Informática de la Universidad Católica de la Santísima Concepción desde 1999. Me doctoré en el Instituto de Investigación en Inteligencia Artificial (IIIA) del Consejo de

Investigaciones Científicas de España en el año 2008 con el apoyo de una beca Conicyt para Doctorado en el extranjero. Mi área de investigación principal es Búsqueda Heurística. La Búsqueda Heurística es un método general de resolución de problemas en Inteligencia Artificial. El tema de mi tesis doctoral fue Búsqueda Heurística en Tiempo Real (BHTR) [1]. Las estrategias de BHTR son métodos de búsqueda “online” que permiten resolver problemas de planificación cuando un agente tiene un tiempo limitado para planificar acciones. Una aplicación de los métodos de búsqueda en tiempo real son los juegos de estrategia para computador tipo World of Warcraft o Baldur’s Gate. Por ejemplo, la compañía “Blizzard Entertainment”, creadora de Warcraft, ha establecido que el tiempo en que los personajes del juego planifican sus movimientos no debe superar los tres milisegundos. La investigación en esta área la he realizado principalmente con mi director de tesis, el Dr. Pedro Meseguer científico titular del IIIA. Nuestras publicaciones más importantes son:

- C. Hernández and P. Meseguer, Lookahead, Propagation and Moves in Real-Time Heuristic Search. In online Proceedings of the International Symposium on Combinatorial Search. SOCS 2009.
- C. Hernández and P. Meseguer, Combining Lookahead and Propagation in Real-Time Heuristic Search. In Proceedings of the First International Symposium on Search Techniques in Artificial Intelligence and Robotics. AAAI 2008.
- C. Hernández and P. Meseguer, Improving LRTA*(k). In Proceedings of the 20th International Joint Conference on Artificial Intelligence, IJCAI 2007, Hyderabad, India.
- C. Hernández and P. Meseguer, Propagating Updates in Real-Time Search: HLRTA*(k). In Proceedings of the 11th Conference of the Spanish Association for Artificial Intelligence.



Carlos Hernández junto a su grupo de investigación.

CAEPIA 2005, Santiago de Compostela, España. Pag. 193-202. LNAI. ISI. Best Paper Award.

- C. Hernández and P. Meseguer, LRTA*(k). In Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI 2005, Edinburgh, Scotland. Pag. 1238-1243.

Desde finales de 2009 he comenzado a trabajar con el Dr. Jorge Baier del Departamento de Ciencias de la Computación de la Pontificia Universidad Católica de Chile en BHTR. Nos aceptaron el trabajo titulado "Escaping Heuristic Hollows in Real-Time Search without Learning" en el Workshop en Agentes y Sistemas Colaborativos (WASC 2010) de las Jornadas Chilenas de Computación 2010, y estamos preparando un artículo para enviar a AAMAS 2011.

Otra línea de investigación que trabajamos desde mediados de 2009 es Búsqueda Heurística Incremental (BHI). La BHI permite resolver problemas de planificación en entornos dinámicos y/o semiconocidos de manera eficiente. El algoritmo más popular del área es D*Lite [2]. Versiones de este algoritmo han sido implementadas en los exploradores de Marte "Spirit" y "Opportunity", y en el automóvil autónomo ganador de la DARPA Urban Challenge - 2009 de CMU. La investigación en BHI la hemos realizado con el grupo de investigación del Dr. Sven Koenig de la University of Southern California¹ y con el Dr. Pedro Meseguer. Hemos publicado el artículo:

- C. Hernández, P. Meseguer, X. Sun and S. Koenig, Path-Adaptive A* for Incremental Heuristic Search in Unknown Terrain. In Proceedings of the 19th International Conference on Automated Planning and Scheduling. ICAPS 2009.

Actualmente trabajamos en un algoritmo que experimentalmente ha mostrado mejoras sustanciales sobre D*Lite. Pretendemos publicar nuestro trabajo en AAMAS-2011 e IJCAI-2011.

La investigación en Búsqueda Heurística que hemos realizado en los últimos años ha sido parcialmente financiada por el proyecto Fondecyt de iniciación Código 11080063 que termina este año. Hemos enviado la postulación al Concurso Regular Fondecyt 2011 en las áreas de Búsqueda Heurística y Planificación Automática en conjunto con el Dr. Jorge Baier de la Pontificia Universidad Católica y el Dr. Mario Medina de la Universidad de Concepción (UdeC).

Desde el año 2009 coordino el proyecto "Técnicas de inteligencia artificial para el uso eficiente del agua en plantaciones de arándanos utilizando redes de sensores inalámbricos²". Éste es financiado por la Fundación para la Innovación Agraria (FIA) del Ministerio de Agricultura. Esta temporada de riego 2010-2011 corresponde implementar el prototipo de nuestro sistema en los huertos comerciales asociados. El trabajo es realizado por académicos y personal contratado por nuestro Departamento. Hemos contado con la colaboración de académicos de la Facultad de Ingeniería

Agrícola de la UdeC y de investigadores del INIA. Pensamos continuar la colaboración en proyectos tecnológicos para el área agrícola postulando a fondos nacionales y extranjeros.

Referencias

- [1] R.Korf. Real-time heuristic search, Artificial Intelligence, Vol. 42, No. 2-3, pp. 189-211, March 1990.
- [2] S. Koenig and M. Likhachev. D* Lite. In Proceedings of the AAAI Conference of Artificial Intelligence (AAAI), 476-483, 2002.

INCA: INVESTIGACIÓN EN INTELIGENCIA COMPUTACIONAL Y APRENDIZAJE AUTOMÁTICO

Departamento de Informática, Universidad Técnica Federico Santa María.

El grupo de Inteligencia Computacional Aplicada (INCA) del Departamento de Informática de la Universidad Técnica Federico Santa María fue formado en el año 2000 por investigadores de las áreas de Estadística Computacional y Aprendizaje Automático. A la fecha, el grupo de investigación es liderado por el profesor titular Dr. Héctor Allende y por el profesor emérito Dr. Claudio Moraga. Además, colaboran en la investigación el Dr. Rodrigo Salas y el Dr.(c) Ricardo Ñanculef; los alumnos de doctorado Carlos Valle, Rodrigo Alfaro, Héctor Allende-Cid, Juan Zamora y Sergio Campos, y el asistente científico Ing. Alejandro Veloz.

El interés de investigación del grupo está centrado en el aprendizaje automático de datos y sus aplicaciones en problemas de pronóstico, procesamiento de imágenes, reconocimiento de patrones y Web data mining, entre otras (ver por ejemplo [1, 9, 12, 13]). Además, se mantienen estrechos vínculos con investigadores de diferentes instituciones europeas, citándose entre ellas: El European Centre for Soft-Computing en

España; el Departamento de Computación del Imperial College of London en Reino Unido; el Departamento de Informática de la Technische Universität Dortmund en Alemania; y el Departamento de Electrónica, Informática y Sistemática de la Università di Bologna en Italia; entre otros centros de investigación.

Sin duda que los progresos conseguidos en el área de aprendizaje automático de datos están cambiando la forma de usar computadores para explorar datos en áreas que se caracterizan por disponer de bases de datos de gran tamaño y alta dimensionalidad, donde se requiere develar patrones y tendencias ocultas demasiado sutiles o complejas para ser detectadas por los seres humanos. Actualmente, existen numerosas aplicaciones de la vida cotidiana donde se están empleando algoritmos de aprendizaje de datos registrados en tiempo real, por ejemplo, en el modelado de varias actividades personales, conversaciones y movimientos, guías de tráfico, entre muchas otras áreas.

El paradigma hoy dominante para construir e implementar algoritmos de aprendizaje automático a partir de un conjunto de datos supone la disposición completa de dichos datos desde el inicio y durante todo el proceso de entrenamiento, de modo que la máquina tiene acceso a todos ellos tantas veces como sea necesario. Sin embargo, existen muchas aplicaciones en las cuales este paradigma deja de ser razonablemente útil y se requiere minimizar por razones prácticas o de factibilidad física los requerimientos de memoria y overhead computacional. Por otro lado, la naturaleza inherentemente dinámica de los fenómenos que generan los datos causa que los patrones característicos estén constantemente evolucionando y cambiando, pudiendo generarse la situación en la cual nuevos conceptos son agregados a un conjunto de datos.

Desafortunadamente, la mayoría de las técnicas de aprendizaje automático están basadas en el supuesto de que los datos se ajustan a una distribución de probabilidad



Grupo INCA.

estacionaria, y se omite la posibilidad de que las distribuciones subyacentes cambien. Este hecho podría degradar el desempeño predictivo de los modelos de máquinas de aprendizaje.

El grupo INCA está abocado principalmente al diseño de algoritmos de aprendizaje automático capaces de extraer y mantener un modelo mediante arribos secuenciales de datos, es decir, los conjuntos de ejemplos en un instante determinado no están disponibles en rondas sucesivas. En este sentido, como parte de nuestro quehacer, hemos explorado y desarrollado metodologías de análisis para los aspectos centrales de los algoritmos de aprendizaje automático, como los son, la capacidad de generalización en ambientes no-estacionarios y dinámicos, y la robustez frente a las desviaciones en los supuestos idealizados de la realidad. Entre los principales resultados que el grupo ha alcanzado está la construcción de nuevas técnicas de procesamiento de datos en base a modelos de máquinas de aprendizaje con arquitecturas flexibles, capaces de adaptarse rápidamente a los ambientes no-estacionarios, pero que al mismo tiempo sean lo suficientemente estables para no verse afectados por la presencia de datos

atípicos o aberrantes en el conjunto de datos. Los principales resultados en este tema se pueden ver en [3, 7, 10, 11].

Las aplicaciones del paradigma de aprendizaje desarrollado por el grupo INCA, se podrían enmarcar en dos escenarios fundamentales. El primero corresponde a aplicaciones de minería de datos con almacenamiento distribuido, en el cual los datos que el algoritmo requiere procesar se encuentran horizontalmente fragmentados sobre una red de nodos de información, a los que el algoritmo debe acceder minimizando la cantidad de datos que se transfieren sobre la red. El segundo escenario corresponde a aplicaciones de minería de datos sobre streams de datos, es decir, a partir de secuencias de ejemplos que arriban a un sistema de manera continua en el tiempo, en volúmenes que no es posible almacenar por completo antes de iniciar el proceso de análisis. Mayores detalles de estas aplicaciones se pueden encontrar en [2, 4, 8].

Nuestros últimos esfuerzos se han concentrado en la formulación de modelos y algoritmos de aprendizaje incremental dentro del marco teórico de las máquinas de soporte vectorial (SVMs, support vector

machines). Específicamente hemos propuesto la utilización de algoritmos de geometría computacional desarrollados recientemente para cómputo de medidas de extensión de conjuntos, los cuales se han mostrado equivalentes a los problemas de optimización subyacentes a un gran número de modelos de soporte vectorial. Por ejemplo, construir una SVM en un problema de clasificación binario resulta equivalente a determinar la distancia mínima entre las envolturas convexas correspondientes a ambas clases de ejemplos. Ante esta equivalencia hemos logrado desarrollar un nuevo método para resolver problemas de clasificación con múltiples categorías, que no requieren la generación de múltiples SVMs binarias y pueden ser más eficientes que otros métodos de estas características propuestos en la literatura. Hemos mostrado también que la equivalencia entre un tipo especial de SVM y el problema de determinar la hipersfera contenedora mínima de un conjunto de puntos (en un determinado espacio característico) permite construir algoritmos muy eficientes para resolver el problema de entrenar una SVM cuando los datos se encuentran distribuidos sobre una red. En este momento, nos encontramos demostrando que esta equivalencia permite la construcción de algoritmos para entrenar aproximadamente una SVM sobre un stream de datos, procesando sólo una vez cada registro, y también que es posible acotar el error acumulado del modelo sobre la secuencia de observaciones. Algunos de los resultados en este tópico han sido publicados en [5, 6].

Referencias

- [1] H. Allende, A. Frery, J. Galbiati, L. Pizarro, "M-Estimators with Asymmetric Influence Functions the Distribution GA0 Case", *Journal of Statistical Computation and Simulation*, Vol. 76 N° 11, pp. 941-956, November, 2006.
- [2] R. Nanculef, C. Valle, H. Allende, C. Moraga, "Ensemble Learning with Local Diversity", *Lecture Notes in Computer: Artificial Neural Networks Science*, Vol. 4113, pp. 264-273, 2006, (ICANN-2006), (Grecia).
- [3] R. Salas, H. Allende, S. Moreno, C. Moraga, "Robust and flexible model of Hierarchical

Self Organizing Map for non-stationary environments", *Journal Neuro-Computing*. Vol. 70 pp.2744-2757, 2007.

- [4] R. Nanculef, C. Valle, H. Allende, C. Moraga, "Two bagging algorithms with coupled learners to encourage diversity", *Lecture Notes in computer Science: Advances Intelligent Data Analysis* Vol. 4723 pp. 130-139, 2007, (IDA-2007, Ljubljana Slovenia).
- [5] R. Nanculef, C. Concha, H. Allende, D. Candell. "Multicategory SVMs by Minimizing the Distances Among Convex-Hull Prototypes". *Proceedings Eight International Conference on Hybrid Intelligence Systems*. Ed. *Proceedings CPS IEEE Computer Society*, Vol. 1 pp. 423-428, 2008. (HIS-2008, Barcelona España).
- [6] R. Nanculef, C. Concha, H. Allende, D. Candell, and C. Moraga, "AD-SVMs: A Light Extension of SVMs for Multicategory Classification", *International Journal of Hybrid Intelligence Systems (JHIS)*, Vol. 6 (2) pp. 69-79 2009.
- [7] C. Saavedra, R. Salas, H. Allende, C. Moraga, "Fusion of Topology preserving Neural Networks". In *Lecture Notes in Computer Science* Vol. 5572, pp. 517-524, 2009. (HAIS-2009, Barcelona España).
- [8] Carlos Valle, Francisco Saravia, Héctor Allende, Raúl Monge, César Fernandez, "Parallel Approach for Ensemble Learning with Locally Coupled Neural Networks", aceptado para publicación en *Neural Processing Letters* (2010).
- [9] H. Allende, D. Bravo, E. Canessa, "Robust Design in Multivariate Systems using Genetic Algorithms", *Journal Quality & Quantity* Vol. 44 No.2 pp 315- 332, 2010.
- [10] R. Salas, C. Saavedra, H. Allende, C. Moraga, "Machine Fusion to Enhance the Topology Preservation of Vector Quantization Artificial Neural Networks", aceptado para publicación en *Pattern Recognition Letters*, 2010.
- [11] S. Campos, R. Salas, H. Allende, C. Castro, "Ensemble of local descriptors with topological preservation for iris pattern recognition", aceptado para publicación en *Pattern Recognition Letters*, 2010.
- [12] H. Allende, C. Moraga, R. Nanculef, R. Salas. "Ensembles Methods for Machine Learning". Chapter of book "Pattern Recognition and Machine Vision" in Honor and Memory of Prof. King-Sun Fu, Editor Patrick Shen-Pei Wang, River Publishing Company, Denmark 2010.
- [13] E. Canessa, C. Droop, H. Allende, "An Improved Genetic Algorithm for Robust Design in Multivariate Systems", aceptado para publicación en *Journal Quality & Quantity* Ed. Springer Verlag, to appear 2011.



John Atkinson.

John Atkinson:

PROCESAMIENTO DE LENGUAJE NATURAL EN LA UNIVERSIDAD DE CONCEPCIÓN

Departamento de Ingeniería Informática y Ciencias de la Computación, Universidad de Concepción.

El lenguaje es uno de los medios de comunicación natural más importante entre humanos, ya sea hablado o escrito. Sin embargo, nuestros mecanismos de procesamiento de lenguaje en sus diferentes etapas son muy complejos tanto del punto de vista fisiológico como cognitivo.

Entender dichos mecanismos, desarrollar teorías, y modelos de procesamiento automático del lenguaje para aplicarlos en la resolución de problemas complejos ha sido foco de investigación mundial desde los inicios de la historia de las Ciencias de la Computación. Por ejemplo, actualmente no sería posible diseñar compiladores para lenguajes de programación si no fuera por los estudios originales de procesamiento de lenguaje del lingüista Noam Chomsky del MIT. Más aún, el test de Turing trabaja sobre la base de interacción humano-computador vía lenguaje natural donde la "máquina" tiene que ser lo bastante inteligente procesando, entendiendo el lenguaje y generando respuestas como para "engañar" a un ser humano.

En la actualidad, existen muchas tareas complejas que involucran el Procesamiento (automático) del Lenguaje Natural (PLN), las que serían prácticamente imposibles de ser

llevadas a cabo por seres humanos. Así, PLN se ocupa de la formulación e investigación de mecanismos computacionalmente eficaces para la comunicación entre personas o entre personas y máquinas por medio de lenguajes naturales. Dada la complejidad de resolver problemas de PLN, se requiere investigación multidisciplinaria proveniente de Inteligencia Artificial (IA), Aprendizaje Automático, Lingüística, Representación de Conocimientos, etc.

Aunque comprender automáticamente el lenguaje natural completo está fuera del alcance actual, se han abordado eficientemente tareas específicas de análisis morfológico, análisis sintáctico, análisis semántico y procesamiento de discurso. Se ha avanzado mucho en resolver problemas específicos de interacción humano-computador o de análisis de fuentes de información no estructuradas disponibles en medios masivos electrónicos (papers, documentos Web, emails, etc.). Áreas de trabajo involucran sistemas de pregunta-respuesta sobre la Web (question-answering), sistemas de diálogos automatizados, extracción de información desde bases de datos documentales, traducción automática, generación de lenguaje natural, descubrimiento de patrones desde textos, reconocimiento del habla, etc.

Actualmente la investigación en PLN realizada en la Universidad de Concepción es la más importante del país y está albergada como parte del Laboratorio de Inteligencia Artificial (www.ia.udec.cl). El trabajo involucra tanto investigación científica como aplicada en NLP en áreas tales como:

- Modelos de diálogo humano-computador para filtrado de información en la Web.
- Generación de Lenguaje Natural a partir de descripciones no-lingüísticas (por ejemplo, describir gráficos estadísticos en lenguaje natural para ser comprensibles por humanos).
- Sistemas de pregunta-respuesta para acceder a respuestas directas a preguntas analizando textos disponibles en medios masivos como la Web, Wikipedia, etc.
- Extracción de información desde grandes bases de datos de documentos médicos

para ser traspasada a bases de datos tradicionales.

- Representación de conocimientos y Ontologías.
- Análisis de opiniones desde textos disponibles en medios electrónicos y redes sociales.
- Descubrimiento de patrones desde textos en lenguaje natural para propósitos de toma de decisiones automatizada, clasificación de documentos, etc.
- Rotulación automática de información semántica y de discurso de textos en lenguaje natural (por ejemplo, convertir información de textos en lenguaje natural puro, a información semiestructurada rotulada del tipo XML, etc.).

El grupo de NLP se encuentra muy activo en la comunidad internacional tanto de NLP propiamente tal como de IA en aspectos tanto científicos como tecnológicos. Además, algunas de las investigaciones realizadas ya están siendo transferidas al sector productivo en la forma de aplicaciones o servicios comerciales. Así, los trabajos ya cuentan con registros de propiedad intelectual a nivel nacional como también están inscritos en registros de licencias en el área de PLN administradas por el Centro Alemán de Inteligencia Artificial (DFKI).

El grupo de PLN es activo en varios proyectos nacionales del tipo Fondecyt, de cooperación internacional del tipo Conicyt TIC-AmSud, ECOS e IBM Research Grants, además de tener cooperaciones activas con otros grupos internacionales en Canadá, Estados Unidos, Francia, Reino Unido y Alemania, entre otros.

NLCM: NATURAL LANGUAGE AND COMPUTING MACHINES

Departamento de Ingeniería Informática y Ciencias de la Computación, Universidad de Concepción.

El grupo NLCM (Natural Language and Computing Machines) del Departamento de Ciencias de la Computación de la

Universidad de Concepción tiene como objetivo investigar aspectos del lenguaje natural que puedan ser formalizados e implementados algorítmicamente para estructurar y consultar objetos visuales. Dos de los proyectos de más impacto de este grupo son:

Estructuración, integración e intercambio de datos contenidos en representaciones gráficas (como gráficos de línea, barra, etc.).

En este proyecto, gráficos estadísticos creados en Excel, GNUPlot, Gnumeric, etc. son analizados y provistos de una semántica enriquecida con la que se pueda construir bases de conocimiento gráficas altamente expresivas. Un ejemplo de consulta a estas KBs sería: “Deme un gráfico que muestre una caída fuerte en el empleo, pero una suba moderada en la venta de autos cero kilómetro”. Para esto, palabras como “fuerte” y “moderado” deben tener una semántica fija, los ejes de los gráficos deben ser de la misma clase, o de una superclase; la semántica de los títulos debe estar estructurada (autos nuevos versus sólo autos), etc. Este proyecto se lleva a cabo con financiamiento de IBM.

Estandarización de un lenguaje escrito para comunicar hechos estadísticos que se suceden en gráficos de línea o barras (“Employment plummeted in July”).

Las agencias estadísticas (como el Instituto Nacional de Estadística en Chile) comunican información que se interpreta de una base de datos numérica de distintos factores de la sociedad: población, empleo, ingreso per cápita, venta de vehículos, etc. Esta comunicación se hace a través de reportes escritos que si bien usualmente poseen gráficos, estos últimos están acompañados, y muchas veces explicados en lenguaje natural. Debido a que palabras como “plummet”, “fall”, “skyrocket” tienen tintes semánticos diferentes (después de todo, las primeras dos palabras son especializaciones de “decremento”, mientras que la última de “incremento”) que colorean la oración en la que aparecen. Estos tintes semánticos tienden a influenciar políticas de estado, y su formalización es altamente deseable para minimizar los efectos de la posible ambigüedad. Este proyecto se lleva a cabo con financiamiento de la Dissemination Division of Statistics, Canadá.



Grupo NLCM.

El NLCM es liderado por el Dr. Leo Ferres, profesor asistente del Departamento de Ingeniería Informática y Ciencias de la Computación de la Universidad de Concepción (como entretenimiento, se pide al lector que lo reconozca en la foto por ser el único que no concuerda con la serie). El Dr. Ferres obtuvo su Doctorado (2005) en Carleton University, Ottawa, Canadá en Ciencias Cognitivas (modelamiento de conductas inteligentes), e hizo su Posdoctorado (2004-2006) en Human-Computer Interaction en la misma Universidad, y fue luego Senior Research Scientist en HCI hasta el 2008, año en que asumió funciones en la Universidad de Concepción. El Dr. Ferres ha sido miembro del comité de programa y publica frecuentemente en ACM ASSETS, CHI, ACL ENLG, UIMA Workshop, W4A, SLPAT entre otras, es revisor de ACM TACCESS, y es el general chair de W4A2011, en Hyderabad, India (parte de la WWW Conference). También fue editor invitado de Taylor & Francis Special Issue on Web Accessibility in the Journal New Review of Hypermedia and Multimedia y ha publicado en el Journal of Web Semantics.

Si el lector ya hizo el ejercicio de identificación más arriba, y para su respuesta, los presentes en la foto son, de izquierda a derecha: Rodrigo Parada, Diego Caro, Leo Ferres, José Fuentes, Rodrigo Zúñiga y Patricio Sandaña.

COMET: COMBINATORIAL OPTIMIZATION AND METAHEURISTICS TEAM

Departamento de Informática, Universidad Técnica Federico Santa María.

COMET, es un grupo de investigación del Departamento de Informática de la Universidad Técnica Federico Santa María (UTFSM), liderado por la Dra. María Cristina Riff, dedicado a la resolución de problemas combinatoriales, usando técnicas provenientes de la Investigación de Operaciones, así como el uso de técnicas modernas inspiradas en el uso de heurísticas, aprovechando las capacidades computacionales disponibles actualmente.

El objetivo de nuestro grupo es desarrollar y proponer métodos basados en heurísticas, y metaheurísticas tales como Algoritmos Genéticos, Simulated Annealing, Tabu Search, Colonias de Hormigas, Sistemas Inmunes Artificiales, Hiperheurísticas y otros algoritmos híbridos, y analizar su aplicación para resolver problemas complejos de satisfacción de restricciones y de optimización en dominios finitos.

Dada la naturaleza de las aplicaciones que consideramos, la investigación incluye además aspectos de paralelismo. Nuestro interés en las técnicas está enfocado principalmente en Control de Parámetros, Hiperheurísticas y Técnicas Autoadaptativas.

Problemas de nuestro interés:

- Constraint Satisfaction
- Strip Packing
- Time Tabling
- Vehicle Routing
- Travel Salesman
- Motion Planning
- Travel Tournament
- Cryptography
- Animations in Games

Realizamos trabajos de cooperación Internacional con Carlos Coello (CINVESTAV, México), Bertrand Neveu (Ecole Nationale des Ponts et Chaussées, Paris, Francia), Olivier Marín (Laboratoire d'Informatique de Paris VI, Francia).

Nuestros ex miembros son: Dr. Jorge Maturana actualmente Universidad Austral, Dr. Carlos Grandón Amadeus Francia, Dr. Marcos Zúñiga Universidad Técnica Federico Santa María, Dr. Ignacio Araya Universidad Técnica Federico Santa María, Dra. Giglia Gómez Universidad de Valparaíso, Msc. Pablo Garrido Universidad Saarland Alemania.

Los miembros actuales de COMET son el investigador Dr. Ignacio Araya; la estudiante de Doctorado Elizabeth Montero; los estudiantes de Magíster Leslie Pérez, Agustín Antonissen y Leopoldo Altamirano, y los estudiantes de Ingeniería Civil: Rodrigo Lisperguier, Ignacio Mella, Daniela López, Felipe Ritz, Renato Rivera.

Nuestros artículos han sido publicados en Journals: Information Sciences, Journal of Heuristics, Neural Computing and Applications, European Journal of Operations Research, Engineering Applications of Artificial Intelligence, International Journal on Artificial Intelligence Tools, Intelligent Automation and Soft Computing Journal,

Nuestro grupo ha publicado recientemente en las conferencias: Genetic and Evolutionary Computation Conference (GECCO), IEEE Conference on Evolutionary Computation (CEC), International Conference on Tools for Artificial Intelligence (ICTAI) PPSN, ISMIS, ICARIS, SOCPAR, NABIC, ICAIS, ALIO/EURO.

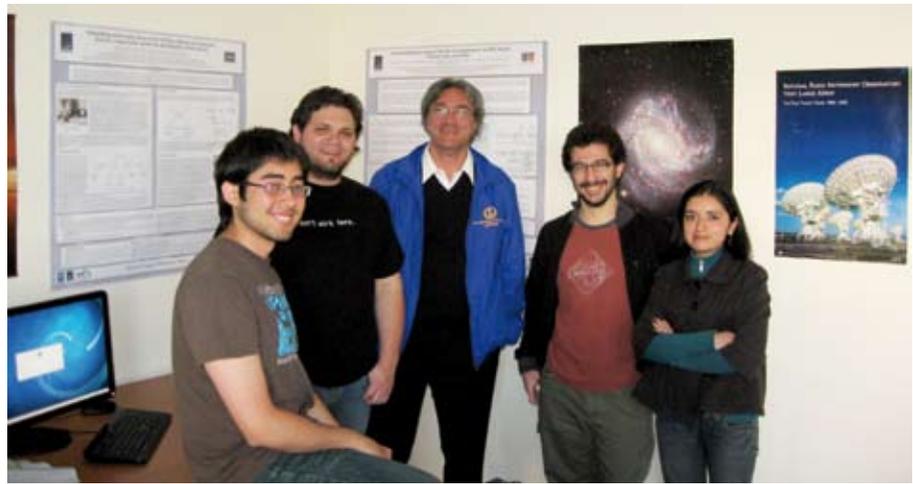
Mauricio Solar:

INTELIGENCIA ARTIFICIAL Y SUS APLICACIONES

Departamento de Informática, Universidad Técnica Federico Santa María.

Planificación y Scheduling son áreas de la Inteligencia Artificial (IA) orientadas a automatizar la ejecución y supervisión de acciones, actividades y/o tareas. El scheduling es un problema NP-completo y se puede resumir como el problema de asignar un conjunto de tareas a un conjunto de recursos sujetos a un conjunto de restricciones. Su implementación proporciona una tecnología que permite (o aumenta) la autonomía de los sistemas con un gran potencial de aplicaciones en diversas áreas.

Es así como se pueden mencionar varias aplicaciones en las que hemos trabajado desde hace más de diez años. Inicialmente buscábamos investigar en nuevos algoritmos de scheduling basados en técnicas heurísticas que han permitido desarrollar herramientas para optimizar el poder computacional que entregan las máquinas con varios procesadores (HPC). De esta forma, los primeros proyectos de investigación apuntaban a desarrollar algoritmos de scheduling para máquinas con memoria compartida (conocidas como de procesamiento simétrico), luego investigando algoritmos para máquinas con memoria distribuida (donde dependiendo de la máquina y los procesos se debe considerar tiempos de comunicación), pasando por la investigación de algoritmos de scheduling para optimizar el uso de los procesadores en un *cluster* (con procesadores homogéneos), hasta el scheduling en computación grid con máquinas heterogéneas, y donde el scheduling no es centralizado sino que se encuentra distribuido en diferentes ambientes. Esta investigación en técnicas heurísticas para desarrollar algoritmos de scheduling condujo la investigación hacia nuevas áreas en las que el scheduling permitiera optimizar el uso de recursos (escasos y caros) dependiendo del área de aplicación. Actualmente, una de las áreas más



Mauricio Solar (al centro) junto a su grupo de investigación.

excitantes en la que estamos trabajando es la investigación de algoritmos de scheduling aplicados a la astronomía.

Scheduling de procesos en máquina paralelas

En esta área hemos investigado algoritmos para asignar código paralelo que debe ejecutarse en máquinas con varios procesadores. Esta investigación se ha financiado con el proyecto Fondecyt N° 1000074 (2000-2002), titulado "Selección Automática de Asignación para Generar Código Paralelo Eficiente", y otro proyecto Fondecyt N° 1030775 (2003-2005), titulado "Desarrollo de Algoritmos de Scheduling para diferentes Modelos de Memoria en Máquinas Paralelas y Distribuidas". Actualmente estamos con un proyecto DGIP-UTFSM (2008-2010), titulado "Desarrollo de Algoritmos Distribuidos de Scheduling para Grid (Grid Scheduling)".

En el contexto de estos proyectos, el trabajo "mdl2dag: Convert Graphical Model System Files to DAGs" (M. Solar y M. Feeley, 2002) presentado en el "Forum on Specification and Design Language" en Marsella (Francia) fue la base para comenzar a trabajar sobre grafos de tareas que representan aplicaciones reales que requerían ser paralelizadas para responder a las exigencias en tiempo real. El alumno de Ingeniería Civil Informática (ICI), Francisco Delgado Álvarez, desarrolló su memoria clasificando estos DAG (Directed Acyclic Graphs) con el objetivo

de posteriormente seleccionar el mejor algoritmo de scheduling para itinerarlo, lo cual fue desarrollado en la tesis de Magíster en Ingeniería Informática (MII) del alumno Mario Inostroza Ponta, dando origen a varias publicaciones entre las cuales se puede mencionar "An Automatic Scheduler for Parallel Machines" (LNCS 2400, pp. 212-216, 2002, M. Solar y M. Inostroza). Esta tesis de Inostroza Ponta permitió generar un código intermedio para multiprocesadores de memoria compartida en la tesis de MII del alumno Héctor Pincheira Conejeros. El alumno Cristián Morales Napoli de ICI de la Universidad Técnica Federico Santa María (UTFSM) implementó un algoritmo genético paralelo para resolver el problema de scheduling en un *cluster* de procesadores homogéneos con excelentes resultados empíricos.

Aplicaciones en Astronomía

En Chile se han construido telescopios de última tecnología y únicos en el mundo. ALMA (Atacama Large Millimeter Array) se está construyendo en la planicie de Chajnantor (cerca de San Pedro de Atacama) a 5000 msnm, producto de una colaboración internacional entre Europa (ESO), EEUU (NRAO) y Asia (NAOJ), que el año 2013 dará origen al radio-telescopio más grande del mundo, con más de 60 antenas. ALMA tendrá la posibilidad de ser usado como un único arreglo, o hasta seis arreglos

independientes o grupos de antenas. Cada arreglo es equivalente a un instrumento, lo cual puede verse como un problema multitelescopio.

En el grupo de CSRG-ALMA de la UTFSM se realizan investigaciones en varias áreas aplicadas de la astronomía. Específicamente en el área de planificación y scheduling en IA, se está desarrollando un proyecto ALMA-Conicyt N° 31080031 (Computer Science for ALMA - Strengthening Research and Development within a Chilean University), en el que se investiga sobre técnicas heurísticas para resolver el problema de scheduling de las observaciones. El objetivo es desarrollar un scheduler que decida qué objeto será observado, cuándo y por cuánto tiempo, optimizando el uso del tiempo de observación y equilibrando consideraciones de corto y largo plazo.

Las preguntas que se busca responder son: ¿Cuándo está disponible el instrumento para los propósitos de cada observación? ¿Cuándo son ideales las condiciones del tiempo para observar un determinado objeto? ¿Cuál objeto debe ser observado primero? ¿Cómo saber cuándo un objeto está listo?

Como primera aproximación en el tema de scheduling en astronomía, el alumno Norman Sáez Vásquez, de ICI de la USM, desarrolló en su memoria un scheduler para el sistema de control de un telescopio amateur, que no tiene las restricciones de un telescopio profesional.

Para construir un scheduling de observación en un telescopio profesional, el sistema debe conocer las condiciones climáticas que cambian dinámicamente, tales como: velocidad del viento y dirección; precipitaciones (nieve o hielo, que pueden detener una observación); humedad; columna de vapor de agua; temperatura; cobertura UV; iluminación del cielo, etc. Otras consideraciones son el número de antenas disponibles en las diferentes configuraciones de ALMA (hasta seis configuraciones), y las condiciones de visibilidad del objetivo, como por ejemplo que se encuentra sobre el horizonte, o que no se encuentre bloqueado por el sol o la luna, etc. Considerando estos requerimientos,

el alumno Arturo Hoffstadt Urrutia de ICI de la USM, desarrolló en su memoria un simulador de las antenas de ALMA para verificar el comportamiento de las antenas en las observaciones astronómicas.

El alumno de MII de la UTFSM, Matías Mora Klein, en su tesis de grado está desarrollando un algoritmo de scheduling con prioridades dinámicas. En un artículo reciente analizó los observatorios y telescopios profesionales más importantes instalados, tales como el telescopio espacial Hubble, Very Large Telescope (VLT), telescopio Subaru, observatorio Gemini, y el telescopio Robert C. Byrd Green Bank (GBT), en el que se concluye que aunque usan cierto grado de scheduling automatizado, todos requieren de intervención de un experto para reconstruir la planificación diaria y tomar decisiones de último minuto.

Como conclusión se puede comentar que el scheduling de las observaciones astronómicas es una nueva instancia del problema de scheduling que es dinámico, grande y complejo, que lo hace fascinante e importante de investigar y resolver.

Acknowledgement

Esta investigación ha sido posible gracias al apoyo financiero del proyecto ALMA-Conicyt #31080031, y AUI.

Jorge Baier, Domingo Mery, Karim Pichara y Álvaro Soto:

GRIMA: GRUPO DE INTELIGENCIA DE MÁQUINA

Departamento de Ciencia de la Computación, Pontificia Universidad Católica de Chile.

En GRIMA nuestro principal objetivo es participar activamente en el desarrollo de nuevas teorías y algoritmos que incrementen el grado de flexibilidad o "inteligencia" de las aplicaciones computacionales actuales. Nuestra visión es que en el corto plazo este tipo de teorías y algoritmos harán posible la aparición de una nueva generación de

aplicaciones computacionales, capaces de operar con alto grado de autonomía en ambientes naturales y virtuales. Posibles escenarios incluyen robots autónomos capaces de operar exitosamente en ambiente naturales no estructurados, o agentes virtuales proactivos capaces de explorar y razonar con diligencia en sistemas de información como la Web.

La creación de este nuevo tipo de aplicaciones impone importantes desafíos, tales como el desarrollo de nuevas representaciones y algoritmos que permitan extraer conocimiento de fuentes de información no estructuradas. Adicionalmente, la operación en ambientes naturales requiere de avanzados sistemas de percepción, tales como algoritmos de visión por computador capaces de realizar tareas como la detección y el reconocimiento de situaciones y objetos. En el caso de sistemas proactivos, es necesaria también la incorporación de algoritmos que permitan la toma de decisiones. En nuestro grupo respondemos a estos desafíos realizando investigación teórica y aplicada en cuatro áreas principales: robótica cognitiva, planeamiento y lógica deductiva, aprendizaje de máquina, y visión por computador. Adicionalmente, nuestras líneas de investigación cubren una serie de temas relacionados como minería de datos, reconocimiento de patrones, sistemas embebidos y procesamiento de imágenes.

En el área de robótica cognitiva hemos desarrollado técnicas para navegación autónoma de robots móviles en ambientes de interior [18,19,20]. Estas técnicas permiten construir mapas de entorno y lograr localización del robot utilizando sensores visuales y de distancia. Las representaciones utilizadas se basan principalmente en grillas de evidencia y mapas topológicos utilizando íconos visuales. En el área de educación hemos creado experiencias que utilizan robots autónomos para apoyar la enseñanza de tópicos como geometría y física a niños en etapa escolar [21,22]. Dada la gran madurez que han alcanzado nuestras técnicas de navegación autónoma y las limitantes que tienen los robots para adquirir información semántica del ambiente,



Karim Pichara, Álvaro Soto, Jorge Baier y Domingo Mery.

actualmente nuestra investigación se ha volcado al desarrollo de nuevos algoritmos de visión por computador para resolver problemas como reconocimiento de objetos, escenas y personas [23,24,25]. Siguiendo las tendencias actuales en el campo de la visión por computador, estos algoritmos están fuertemente basados en el uso de técnicas de aprendizaje de máquina, en particular, modelos probabilísticos. Finalmente, parte importante de nuestra investigación también se centra en el ámbito de planeamiento, en particular, en el desarrollo de técnicas para manejar la típica disyuntiva de un agente autónomo entre disminuir sus incertezas o lograr sus objetivos (exploración versus explotación) [26].

En el área de representación de conocimiento y razonamiento investigamos representaciones lógicas y algoritmos que permiten a un agente inteligente actuar hábilmente en su ambiente. Nuestro foco principal en este último tiempo ha sido el desarrollo de algoritmos eficientes para resolver el problema de planning, el cual consiste en encontrar un curso de acción para un agente dados tres parámetros: un objetivo a cumplir, una representación de la dinámica del mundo y un objetivo. Planning tiene múltiples aplicaciones: desde implementación de robots hasta composición de componentes de software. Específicamente, hemos estudiado cómo es posible resolver el problema en forma efectiva cuando los objetivos son temporalmente extendidos [13] o hay

existencia de preferencias [14] de usuarios. Además, nos interesa especialmente la interacción de nuestras técnicas con las de otras áreas de Ciencia de la Computación. Como ejemplo, hemos investigado cómo las técnicas de planning pueden ser aplicadas a la verificación de software [15] y cómo técnicas de análisis estadístico de lenguaje natural [16] se aplican a planning. Actualmente también trabajamos en aplicaciones de estas técnicas a áreas de ingeniería como el diagnóstico de dispositivos en ambientes dinámicos [17].

En el área de aprendizaje de máquina investigamos en diversas técnicas aplicadas al análisis de información [5,6,12], entre ellas selección de variables, detección de anomalías y aprendizaje activo [1,2,3,4]. La selección de variables es de gran utilidad en la mayoría de los problemas de aprendizaje de máquina, ayuda a mejorar el rendimiento en problemas donde existen muchas variables que pueden confundir los

procesos de inferencia de conocimiento [4]. La detección de anomalías es un desafío que cada vez se hace más relevante [2,5], por ejemplo, hoy son necesarios sistemas de detección de fraudes bancarios, fallas en procesos productivos o detección de objetos nuevos en bases de datos de astronomía. Desarrollar este tipo de sistemas permite obtener información valiosa en casos donde realizar la tarea de detección manual ya no es posible. El aprendizaje activo también constituye un área muy relevante en el desarrollo de las últimas tecnologías sobre el aprendizaje de máquina [2], consta de crear sistemas capaces de elegir las situaciones que utilizará para aprender un modelo en forma semisupervisada. La motivación está basada en la capacidad que tenemos los humanos de retener sólo las situaciones que sabemos son más representativas de un problema en particular y deseamos las situaciones que no nos aportan mayor conocimiento y pueden ser omitidas a la hora de aprender.

En el área de visión por computador hemos desarrollado fuertemente aplicaciones industriales para la caracterización de alimentos y materiales usando imágenes radiográficas y ópticas. En los últimos años hemos participado en proyectos I+D en: inspección de color y forma de filetes de salmón [7], control de calidad de tortillas mexicanas [8], detección de espinas en salmones [9], caracterización de papas fritas [10]. La metodología empleada en estas aplicaciones se basa en un mismo esquema de reconocimiento de patrones, donde se extraen en cada imagen más de dos mil características visuales, de las cuales no más de veinte son usadas por un clasificador. La selección de características y del clasificador se lleva a cabo de manera

	2006	2007	2008	2009	2010	TOTAL
Publicaciones ISI	8	10	7	4	7	65
Pub. En Congresos	9	14	11	12	8	101
Citaciones	43	71	69	52	36	321
Proy. de Investigación	3	6	7	8	6	17
MSc Students	6	5	6	8	11	11
PhD Students	-	5	7	8	10	10

automática en una fase de entrenamiento [11]. Hoy en día, mediante un proyecto Fondecyt, estamos investigando la detección de anomalías internas en objetos usando inspección radiográfica de múltiples vistas con técnicas de view planning y active learning. La idea es contar con un manipulador robótico que pueda ubicar el objeto según lo que indique un algoritmo que, a partir del análisis de las imágenes radiográficas, determinará de manera activa las posiciones de interés a inspeccionar. Una posible aplicación sería la detección de objetos peligrosos en maletas.

En GRIMA hoy contamos con cuatro profesores jornada completa, más de quince cursos en el área, dos estudiantes de doctorado graduados este año y uno por graduarse en los próximos meses. A la vez contamos con cuatro profesores invitados: Miguel Torres y Cristián Tejos del Departamento de Ingeniería Eléctrica de la PUC, Pablo Zegers de la Facultad de Ingeniería de la Universidad de los Andes y Miguel Carrasco de la Escuela de Informática y Telecomunicaciones de la Universidad Diego Portales. En la tabla se puede apreciar algunas estadísticas relevantes.

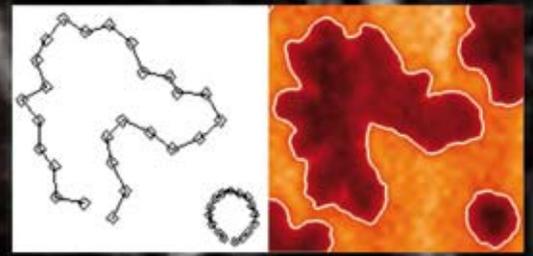
Más información en nuestro sitio Web: grima.ing.puc.cl.

Acknowledgments

This work was partially funded by FONDECYT grants 1095140 and 1100830, Millennium Nucleus for Plant Functional Genomics (P006-09-F), Fondef grant D07I1080, and LACCIR Virtual Institute grant R1208LAC005. BITS

Referencias:

- [1] K. Pichara and A. Soto, "Learning Discriminative Subsets of Features for Classification Using Gaussian Processes". Submitted to International Conference in Data Mining 2010, waiting for revision.
- [2] K. Pichara and A. Soto, "Active Learning and Subspace Clustering for Anomaly Detection". Intelligent Data Analysis (IDA), ISSN: 1088-467X, Volume 15 (2), 2011.
- [3] K. Pichara, A. Soto, and A. Araneda, "Detection of Anomalies in Large Datasets Using an Active Learning Scheme Based on Dirichlet Distributions". Advances in Artificial Intelligence, Iberamia-08, LNCS 5290, pp. 163-172, 2008. Best Student Paper Award.
- [4] T. Puelma, A. Soto, and R. Gutiérrez, "An ensemble of Discriminative Local Subspaces in Microarray Data for Gene Ontology Annotation Predictions". Proc. of 1st Chilean Workshop on Pattern Recognition (CWPR), pp. 52-61, 2009.
- [5] A. Cansado and A. Soto, "Unsupervised Anomaly Detection in Large Databases Using Bayesian Networks". Applied Artificial Intelligence, vol. 22, No. 4, pp. 309 – 330, 2008.
- [6] Urtubia, J. R. Pérez-Correa, A. Soto, and P. Pszczółkowski "Using Data Mining Techniques to Predict Industrial Wine Problem Fermentations". Food Control, vol. 18, No. 12, pp. 1512–1517, 2007.
- [7] Aguilera, J.M.; Cipriano, A.; Eraña, M.; Lillo, I.; Mery, D.; Soto, A.; Valdivieso, C. (2007): Computer Vision for Quality Control in Latin American Food Industry, A Case Study. International Conference on Computer Vision (ICCV-2007): Workshop on Computer Vision Applications for Developing Countries, Rio de Janeiro, Oct. 15.
- [8] Mery, D.; Chanona-Pérez, J.; Soto, A.; Aguilera, J.M.; Cipriano, A.; Velez-Riverab, N.; Arzate-Vazquez, I, Gutiérrez-Lopez, G. (2010): Quality Classification of Corn Tortillas using Computer Vision. Journal of Food Engineering, 101(4):357-364.
- [9] Mery, D.; Lillo, I.; Loebel, H.; Riffo, V.; Soto, A.; Cipriano, A.; Aguilera, J.M.: Automated Detection of Fish Bones in Salmon Fillets using X-ray Testing. In Proceedings of 4th Pacific-Rim Symposium on Image and Video Technology (PSIVT2010), Singapore, Nov.14-17, 2010.
- [10] Pedreschi, F.; Mery, D.; Bungler, A.; Yañez, V.: Computer Vision Classification of Potato Chips by Color. Journal of Food Processing Engineering (accepted Sep 2009).
- [11] Mery, D., Soto, A. (2008): Features: The more the better. The 7th WSEAS International Conference on Signal Processing, Computational Geometry and Artificial Vision (ISCGAV-2008), Rodos Island, Greece, August 20-22.
- [12] A. Soto, F. Zavala, and A. Araneda. "An Accelerated Algorithm for Density Estimation in Large Databases, Using Gaussian Mixtures". Cybernetics and Systems, vol. 38, No. 2, pp. 123-139, 2007.
- [13] Baier, J. A. and McIlraith, S. A. (2006a). Planning with First-Order Temporally Extended Goals Using Heuristic Search. In Proceedings of the 21st National Conference on Artificial Intelligence (AAAI), pages 788-795. Boston, MA.
- [14] Baier, J. A., Bacchus, F., and McIlraith, S. A. (2009). A Heuristic Search Approach to Planning with Temporally Extended Preferences. Artificial Intelligence 173(5-6):593-618.
- [15] Albarghouthi, A., Baier, J. A., and McIlraith, S. A. (2009). On the Use of Planning Technology for Verification. In Proceedings of ICAPS Workshop on Verification and Validation of Planning and Scheduling Systems.
- [16] Muise, C., McIlraith, S., Baier, J. A., and Reimer, M. (2009). Exploiting N-gram Analysis to Predict Operator Sequences. In Proceedings of the 19th International Conference on Automated Planning and Sched. (ICAPS). Thessaloniki, Greece.
- [17] Sohrabi, S., Baier, J., and McIlraith, S. A. (2010). Diagnosis as Planning Revisited. In Proceedings of the 12th International Conference on Knowledge Representation and Reasoning (KR). Toronto, Canada.
- [18] A. Araneda, S. Fienberg, and A. Soto, "A Statistical approach to simultaneous mapping and localization for mobile robots". The Annals of Applied Statistics, vol. 1, No. 1, pp. 66-84, 2007.
- [19] P. Espinace, D. Langdon, and A. Soto, "Unsupervised Identification of Useful Visual Landmarks Using Multiple Segmentations and Top-Down Feedback". Robotics and Autonomous Systems, vol. 56, No. 6, pp. 538-548, 2008.
- [20] P. Espinace, A. Soto, and M. Torres-Torriti, "Real-Time Robot Localization In Indoor Environments Using Structural Information". IEEE Latin American Robotics Symposium (LARS), 2008.
- [21] R. Mitnik, M. Recabarren, M. Nussbaum, and A. Soto, "Collaborative Robotic Instruction: A Graph Teaching Experience", Computers & Education, vol. 53 , No. 2, pp. 330-342, 2009.
- [22] R. Mitnik, M. Nussbaum, and A. Soto, "An autonomous educational mobile robot mediator". Autonomous Robots, vol. 25, No. 4, pp 367-382, 2008.
- [23] P. Espinace, T. Kollar, A. Soto, and N. Roy, "Indoor Scene Recognition Through Object Detection". In Proc. of IEEE International Conference on Robotics and Automation (ICRA-2010).
- [24] S. Montabone and A. Soto, "Human Detection Using a Mobile Platform and Novel Features Derived From a Visual Saliency Mechanism". Image and Vision Computing, vol. 28, No. 3, pp. 391-402, 2010.
- [25] D. Maturana, D. Mery, and A. Soto: Face Recognition with Decision Tree-based Local Binary Patterns. In Proceedings of Asian Conference on Computer Vision (ACCV2010), Queenstown, Nov.08-12, 2010.
- [26] J. Correa and A. Soto, "Active visual perception for mobile robot localization". Journal of Intelligent and Robotic Systems, vol. 58, No. 3-4, 2010, pp. 339-354.



Contornos activos para segmentación en imágenes digitales

Contornos activos sobre imagen de capas de lípidos.

Los continuos avances y la masificación de la computación y la electrónica en décadas recientes han repercutido con fuerza en el desarrollo del procesamiento de imágenes y su impacto en diversos ámbitos del quehacer humano, aprovechando la información que es posible obtener a partir de una imagen. Sistemas satelitales, microscopía, cámaras de todo tipo, ultrasonido, resonancias magnéticas, entre muchas otras fuentes, generan un volumen ingente y en constante aumento de datos que sirven a procesos de registro, inspección y/o control visual, con propósitos que varían según el área de aplicación y aun con los objetivos o escenarios particulares de sus usuarios: robótica, entretenimiento, investigación científica, medicina, vigilancia y procesos de manufactura, por nombrar algunos.

Sistemas como procesadores digitales de señales con capacidades en constante mejora, métodos como agrupamiento (*clustering*) o reconocimiento de patrones, y

los sistemas de cómputo de uso general han permitido la aplicación de métodos de cada vez más alto nivel para abordar problemas de complejidad y tamaño creciente en el tiempo. Por ejemplo, los microscopios ópticos modernos pueden recolectar cientos de gigabytes en minutos para estudiar procesos biológicos in vivo, con decenas de imágenes por segundo; telescopios de última generación pueden producir imágenes ópticas o de radiofrecuencia del cielo, del orden de terabytes por minuto (como los proyectos en desarrollo para telescopios de radio y ópticos en el norte de Chile).

Tanto el volumen como la complejidad de los datos hacen necesario disponer de herramientas y modelos para su almacenamiento, tratamiento y análisis con algún nivel de automatización, constituyendo actualmente desafíos para matemáticas, computación, electrónica y disciplinas afines. A continuación se presentan los conceptos generales de los modelos de



Jorge Jara

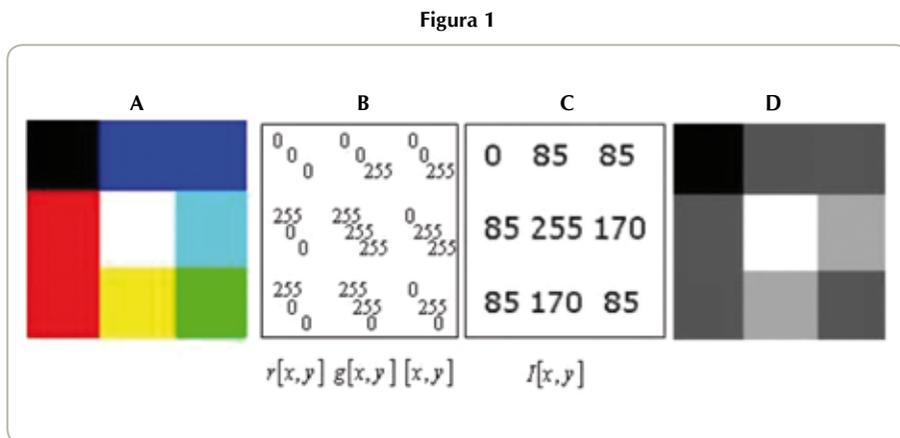
Estudiante, Programa de Doctorado en Ciencias mención Computación DCC, Universidad de Chile. Ingeniero Civil en Informática, Licenciado en Cs. de la Ingeniería mención Informática, Universidad Austral de Chile. Líneas de investigación: Procesamiento de Imágenes y Segmentación, Geometría Computacional.
jjara@dcc.uchile.cl

contorno activo, que son métodos para modelar objetos de interés en imágenes, y que se encuentran dentro de las llamadas técnicas de segmentación.

ALGUNOS CONCEPTOS PREVIOS

Consideremos una *imagen* como una representación de uno o más objetos, generada por proyecciones basadas en las variaciones de la luz sobre ellos. En forma más general, es posible generar imágenes a partir de ultrasonido, resonancia magnética, temperatura u ondas de radio. Cualquier imagen en n dimensiones puede ser expresada como una función $I = I(x_1, x_2, \dots, x_n)$, cuyos valores son una característica observable o medible como intensidad de luz o color, profundidad, temperatura, densidad, etc. (un ejemplo muy simple se muestra en la fig. 1). Por ejemplo, una fotografía digital es una función de dos dimensiones (un plano) que podemos ver como una cuadrícula, en cuyos casilleros (píxeles) se registra el color de la escena fotografiada: la cámara que registra la imagen cuantifica el color de la escena y lo codifica con un valor numérico en cada píxel. Dispositivos más sofisticados son capaces de registrar imágenes tridimensionales de objetos en aplicaciones que van desde la microscopía atómica hasta vastas regiones del universo.

El *procesamiento de imágenes* se entiende como uno o más procesos de manipulación y/o tratamiento sobre un conjunto de imágenes de entrada para generar una



Formación de una imagen digital. Una imagen digital en escala de grises se define con una sola matriz $I[x, y]$. Cada posición de la matriz define un píxel y tiene valores en una escala que representan la intensidad I ; en escala de 8 bits, los valores entre 0 y $2^8 - 1 = 255$ codifican desde el negro hasta el blanco. Una forma para presentar colores en imágenes digitales imita al sistema visual humano: una imagen en color se puede definir combinando tres matrices o canales: $r[x, y]$ para el canal rojo (red), $g[x, y]$ para el canal verde (green), y $b[x, y]$ para el canal azul (blue). A: imagen digital en colores. B: representación r-g-b de los valores de intensidad para la imagen en A. C: imagen en escala de grises, con un solo valor de intensidad asociado (D). D: codificación en escala de grises. Fuente: www.scian.cl.

cierta salida. La diversidad y cantidad de procesos es tan amplia que se han distinguido tipos de procesos de acuerdo a su finalidad. Distingamos los siguientes tipos de procesamiento:

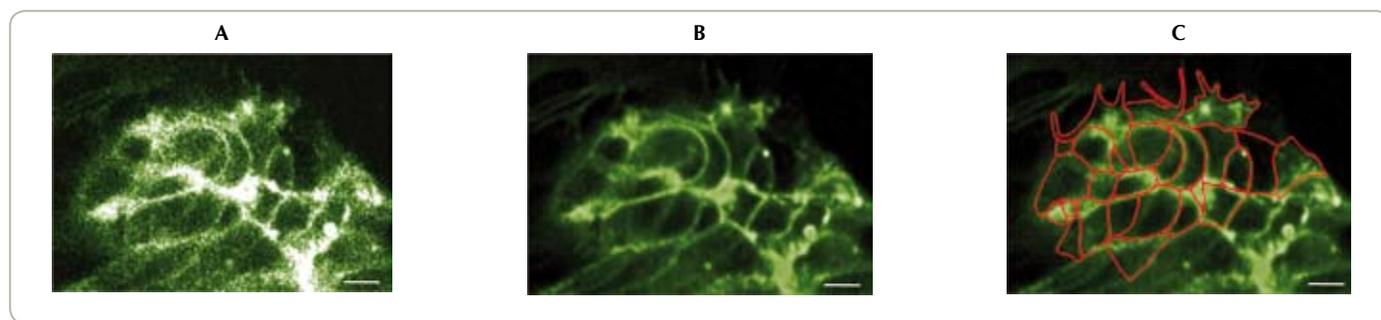
- Tratamiento de imágenes, que apunta a recrear una imagen lo más fiel posible a la realidad que representa, quitando o minimizando la influencia de elementos ajenos introducidos en la adquisición o almacenamiento de la imagen (como baja iluminación o ruido, por ejemplo).
- Análisis de imágenes, busca generar descripciones sobre una o más imágenes dadas, que den cuenta de características de interés.

- Comprensión de imágenes, a un nivel de abstracción mayor, busca interpretar y describir una o más imágenes en el contexto del problema que lo requiere.

Aunque esta clasificación es bastante general y algo simplificada, y existen otras que no necesariamente coinciden con ella, servirá como marco general para contextualizar el problema de segmentación como parte de una problemática más amplia y compleja.

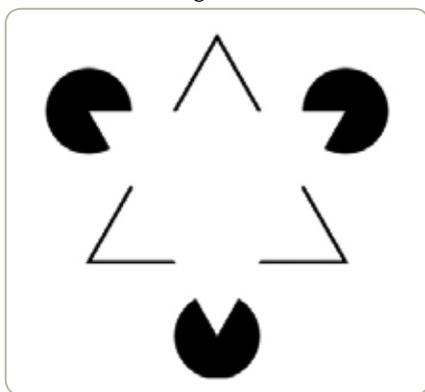
La fig. 2 muestra una imagen digital de células observadas mediante microscopía, que pasa por una etapa de tratamiento para minimizar el ruido y la distorsión del

Figura 2



Tratamiento y segmentación (análisis) en una imagen digital. A: imagen de microscopía de un conjunto de células cerebrales en un embrión de pez cebra, obtenidas mediante escáner de barrido láser. B: imagen tratada para minimizar la distorsión del proceso de adquisición empleando algoritmos de propósito específico. C: segmentación de bordes de estructuras celulares sobre la imagen; cada borde da cuenta de una región de interés, y su representación computacional permite acceso a descripciones geométricas de características como tamaño y forma. La barra de escala corresponde a $5 \mu\text{m}$. Fuente: SCIAN-Lab (datos no publicados).

Figura 3



Ejemplo de imagen con figuras subjetivas, los triángulos de Kanizsa.

microscopio, y que en una etapa de análisis es “descrita” mediante modelos de curva cerrada que demarcan las membranas en cada célula.

SEGMENTACIÓN

La segmentación es el proceso de subdividir una imagen en un cierto número de regiones, teniendo cada una de ellas alguna propiedad distintiva. Cuando se busca distinguir a un conjunto de objetos que presentan una propiedad específica, dichos objetos son denominados regiones de interés (*regions of interest, ROIs*): las personas en una fotografía o estrellas en una imagen de telescopio son ejemplos de ROIs. En general, las regiones son conjuntos disjuntos en la imagen, es decir que no se sobrelapan, de modo que cada elemento de la imagen (por ejemplo, un píxel en una imagen digital) pertenece a una sola región (se pueden hacer excepciones según el problema, cuando objetos se superponen o “comparten” una región del espacio, o bien no es posible resolver la imagen con más detalle para separarlos). El propósito de la segmentación es facilitar el acceso a información relativa a la imagen o a sus elementos constituyentes para etapas posteriores de análisis y extracción de información. El primer problema es que la definición misma de cada región es un problema que involucra aspectos psicológicos y físicos, (ver por ejemplo la fig. 3), de modo que no es posible establecer un criterio único u objetivo como *gold standard* para evaluar la pertinencia o calidad de una u otra técnica

de segmentación. La variedad de técnicas y formas de implementación existentes a la fecha, en combinación con lo anterior, obstaculizan la existencia de un enfoque unificado de segmentación, a pesar de lo cual existen trabajos en el área que apuntan a definir algunos criterios para cuantificar la calidad de distintos métodos.

La segmentación aparece como uno de los desafíos actuales en el procesamiento de imágenes a gran escala, como ocurre en biología y astronomía, en atención a requerimientos como la precisión y gran volumen de datos, objetos de morfología compleja y características particulares (por ejemplo, la fig. 2 que muestra entramados celulares y membranas con pequeñas prolongaciones a escala micrométrica). Métodos de segmentación han sido motivados específicamente por el análisis de imágenes en un campo particular, o suelen ser objeto de análisis y aplicación de técnicas nuevas, tanto por la masificación de tecnologías de computación como por sistemas ópticos que permiten observaciones a escala micro y nanométrica, muy cercana al nivel molecular capturando decenas o cientos de gigabytes de imágenes por experimento. Además, el análisis de objetos y complejos que cambian en el tiempo requiere tanto de modelos como implementaciones apropiadas y eficientes para su identificación y posterior descripción.

¿Cómo segmentar?

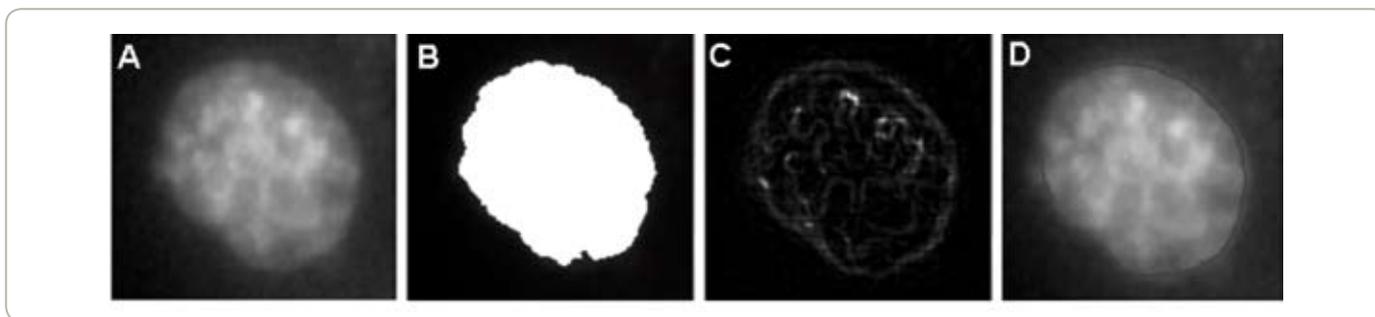
Con la premisa de que es de interés encontrar y describir ciertos objetos o regiones en una imagen, se puede aprovechar la forma en que la representamos: al considerarla como una función matemática, pueden evaluarse propiedades o características que permiten identificar o entregar pistas sobre las regiones de interés. Una forma directa es detectar **los cambios de color o intensidad** que existen en una imagen: la idea es que si cada objeto presenta un patrón de color o brillo relativamente homogéneo, se puede cuantificar la diferencia que se produce cuando se cambia de una región a otra (ver ejemplo en la fig. 4). En escenarios más complejos se pueden ajustar plantillas o patrones de formas, como por ejemplo probar

en qué sitio y de qué forma encaja mejor un rectángulo, una silueta de persona, etc. También se puede aprovechar información extra proporcionada junto con la imagen, como número de objetos, tamaños u otras características que “guíen” a la segmentación. Aprovechando las particularidades de distintas clases de imágenes, existe una gran variedad de modelos matemáticos y computacionales para segmentación, con diferentes grados de automatización y complejidad; ejemplos de esto lo constituyen modelos de *pattern matching*, *clustering*, redes neuronales, probabilísticos y funciones de optimización.

MODELOS DE CONTORNO ACTIVO

Los contornos activos son modelos de optimización que buscan balancear propiedades de imagen que definen a cada ROI junto con características del contorno de cada región (como por ejemplo su regularidad o curvatura), que son definidas a priori al formular el modelo de optimización: el contorno de cada región es modelado como una estructura elástica sometida a fuerzas que la deforman hasta que alcanza un estado de equilibrio – el óptimo – con mínima energía; por esta razón también se habla de “modelos deformables”. El estado óptimo se determina comúnmente mediante ecuaciones diferenciales que definen una condición de equilibrio entre distintas propiedades o *fuerzas*: por un lado se definen las fuerzas internas o propiedades intrínsecas a la forma del contorno, mientras que las fuerzas externas son las características de la imagen (como las transiciones de color o intensidad); también es posible definir fuerzas que mejoran la convergencia del ajuste o permiten completar información insuficiente en la imagen. La gran ventaja de los modelos de contorno activo es que admiten una amplia gama de formas, puesto que no condicionan una representación geométrica específica, sino que involucran varios grados de libertad que se ven restringidos por principios físicos que determinan su comportamiento según las propiedades definidas para las regiones a segmentar.

Figura 4



Segmentaciones por umbral y contornos activos en una imagen de escala de grises. A: imagen en escala de grises de un núcleo celular, obtenida por microscopía. B: Segmentación del núcleo (región de interés, en blanco) utilizando un umbral de intensidad, todos los píxeles con una intensidad mayor o igual al valor umbral son considerados parte del núcleo. C: imagen de gradientes de intensidad; se calcula la magnitud de los cambios de intensidad en la imagen original (A), en que los píxeles más brillantes corresponden a mayores gradientes o “saltos” de intensidad. D: segmentación del núcleo mediante contorno activo, utilizando como base el contorno de la región segmentada en B, y como fuerza de atracción la imagen de gradientes de C. Fuente: SCIAN-Lab (datos no publicados).

Los modelos de contorno han sido incluidos en formulaciones más generales, que permiten considerar otros elementos tales como las propiedades de la imagen que queda dentro y fuera de cada ROI, o el traslape de objetos. La extensión y complejidad del tema escapa al alcance de este artículo, por lo que se remite al lector interesado a la bibliografía.

Se distingue entre las formulaciones explícita e implícita, según la función de contorno, siendo representativos de cada una los snakes y los contornos activos basados en *level sets*, respectivamente, que se presentarán a continuación.

Si bien hasta ahora el modelamiento aparece más como un problema de corte más matemático que de “computines”, la conjugación matemático-computacional se da actualmente en equipos de trabajo y laboratorios de investigación de muchos países, con aplicaciones en robótica, astronomía, biología y medicina, por nombrar algunas. ¿La razón? Necesitamos adentrarnos un poco más en el tema...

MODELOS EXPLÍCITOS O PARAMÉTRICOS

Kass, Witkin y Terzopoulos [2] presentaron en 1988 un modelo para 2D que llamaron *snake* (serpiente): una curva en un plano es deformada por ecuaciones que pesan fuerzas internas y externas. Como fuerzas

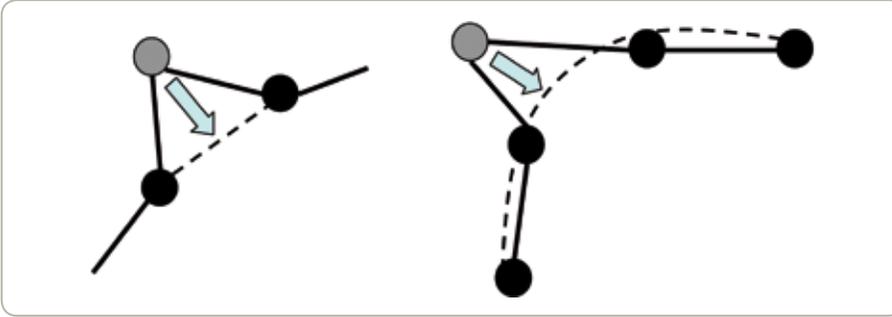
internas definieron elasticidad y rigidez, y como fuerza externa gradientes de intensidad de la imagen (fig. 5). Modelaron la curva como un conjunto finito de puntos $[x, y]$, representable en un programa de computador, con varios pasos de deformación hasta alcanzar un estado de reposo (óptimo), que definieron como el mínimo valor de la suma de las fuerzas evaluadas sobre todo el contorno, en forma de integral. La solución a esta minimización de la integral se puede obtener mediante ecuaciones diferenciales, definiendo ciertas restricciones y calculando las fuerzas en cada punto del contorno, deformándolo en iteraciones sucesivas desde un estado inicial. El estado inicial podía ser definido por el usuario, dibujando algunos puntos de control que permitían generar una interpolación para completar la curva. Además, Kass y cols. incluyeron fuerzas de atracción y repulsión definibles por el usuario, en forma de puntos específicos que llamaron de *resorte* y *volcán*, respectivamente; de este modo podían evitar que el *snake* se deslice hacia zonas no deseadas en la imagen.

MODELOS IMPLÍCITOS O GEOMÉTRICOS

El mismo año 1988, Osher y Sethian [3] presentaban un trabajo de física, en que modelaban la propagación de frentes; un anillo de fuego que se expande consumiendo un pastizal es un ejemplo. La expansión de

frente se modeló utilizando la curvatura y un tipo de ecuación que se popularizó por sus numerosas aplicaciones, una de las cuales se presentó más tarde como un modelo geométrico para contornos activos [4]. La idea general es que una curva cerrada en el plano se puede deformar con velocidad proporcional a su curvatura (mientras más plana se mueve más rápido, por ejemplo): si a esto se agregan fuerzas de imagen se obtiene un modelo parecido al de los *snakes*. Lo novedoso para este caso es el método que se emplea para resolver la ecuación: imaginemos primero que la curva de contorno 2D es parte de una superficie 3D que se deforma siguiendo una ecuación que incluye las reglas de evolución del contorno 2D; ahora pensemos que esta superficie es una especie de “mapa de elevación” en que la altura cero coincide con el contorno de la región, que se llama “curva de nivel cero” (de ahí el nombre de *level sets*, ver fig. 6); esta superficie se deforma siguiendo la ecuación de ajuste hasta que alcanza un estado de equilibrio, y se rescata el resultado buscando las curvas de nivel cero sobre la imagen. Este método permite que el número de contornos obtenidos sea independiente de con cuántas curvas empezó el ajuste. Mientras en el modelo de los *snakes* el número de ROIs debe coincidir con el número de curvas iniciales, para los *level sets* esto no es un problema, incluso es posible demostrar que su resultado (óptimo) no depende de cómo se inicialice el método.

Figura 5



Elasticidad y rigidez en un *snake*. En una curva de contorno, representada por un número finito de puntos, las propiedades o fuerzas de deformación (definidas en forma analítica) son calculadas aplicando una versión aproximada para cada punto, en función de sus puntos vecinos. Se muestra el efecto de calcular dichas fuerzas en un punto de prueba (gris). Izquierda: el efecto de fuerza elástica mueve al punto de prueba hacia la posición “promedio” entre sus vecinos. Derecha: el efecto de la fuerza de rigidez mueve al punto de prueba hacia la posición que corresponde a la curvatura estimada en sus vecinos.

EL ROL DE LA COMPUTACIÓN

El problema de contornos activos es un problema de optimización, y como tal se espera que su solución exista y se pueda aproximar por algún método, y es aquí donde entra la computación en juego. En la actualidad hay líneas de investigación activas en matemáticas al respecto, que buscan garantizar las condiciones de solución y convergencia de los métodos, así como formular propiedades que permitan segmentar imágenes con ruido, información faltante, oclusión de objetos, regiones con distintas propiedades, etc. Por su parte, el modelamiento y la implementación computacional en imágenes digitales que se valen de recursos algorítmicos, ha permitido aplicar diversos modelos de contorno activo con éxito en varias aplicaciones. A continuación veamos algunos ejemplos en que el uso de técnicas de algoritmos y geometría permite superar limitaciones de un modelo en el sentido matemático:

Snakes adaptativos. El método de los *snakes* es popular porque se formula y entiende con relativa facilidad, a pesar de que requiere de una buena inicialización, y la falta de una formulación que garantice una solución única. Para ambos casos han surgido algoritmos que permiten superar este tipo de obstáculos:

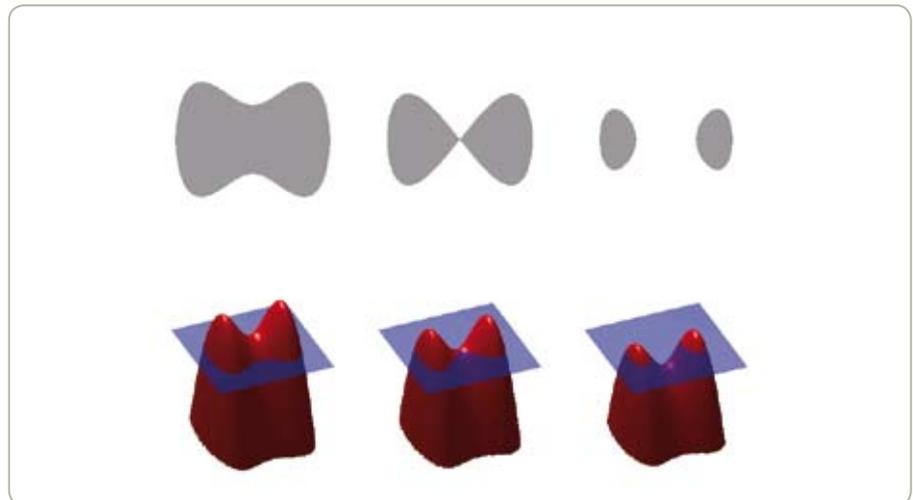
1. Si se inicializa una gran cantidad de contornos sobre la imagen y se sigue su evolución, se puede controlar cómo fusionar o dividir curvas: en el método de los *topology adaptive snakes* o *t-snakes* (publicado oficialmente años después de ser presentado en conferencia, [5]) se coloca una grilla sobre la imagen, que permite ir chequeando los cambios en el tiempo de los distintos *snakes*, y decidir si se eliminan, fusionan o separan curvas. De este modo es

posible utilizar una inicialización arbitraria, permitiendo incluso realizar segmentaciones automáticas cuando los parámetros del modelo son compatibles con las imágenes de entrada. A la fecha existen implementaciones tanto en 2D como en 3D.

2. La dependencia de la inicialización significa que un *snake* puede quedar “atrapado” en una zona de la imagen en que la suma de fuerzas sea cero pero que no corresponda a los bordes de una ROI, lo que vendría a corresponder con un óptimo local del problema de optimización. Cohen [6] presentó las fuerzas de inflación (que llamó *balloon forces*) con la siguiente idea: si un *snake* se encuentra en una zona que se considera interior de la ROI, se expande o “infla” (deformación en dirección normal hacia fuera), mientras que si esto ocurre en una zona considerada como exterior en la imagen el *snake* se contrae o “desinfla”.

Acelerando el cómputo de los level sets. Las ecuaciones de *level sets* se resuelven sobre la imagen completa como dominio, calculando una superficie iterativamente hasta que el resultado converge, para luego

Figura 6



Un ejemplo de curvas de nivel o *level sets* para contornos de formas en 2D. Arriba: tres regiones 2D en gris. Abajo: funciones de *level sets* de las regiones 2D; para cada región se tiene una función de superficie 3D, ϕ , cuyo valor cero (el contorno de cada región 2D) se muestra como la intersección con el plano de la imagen. Fuente: archivo de Wikimedia Commons, con licencia de dominio público.

realizar un posproceso de recuperación de las curvas de nivel. Pensemos en una imagen 2D compuesta por n píxeles en total. Con los métodos tradicionales de diferencias finitas o elementos finitos podemos requerir $O(n^3)$ operaciones para obtener la solución. Si los contornos (curvas de nivel cero) son muy pocos y/o pequeños hemos desperdiciado espacio y tiempo en el cálculo de porciones de la solución que no son de utilidad. Esta idea motiva el uso de algoritmos de cálculo de bandas (*narrowband level sets*), en que sólo se calcula la solución para una vecindad del nivel cero, la que se actualiza con cada iteración; también es posible hacer uso de técnicas de compresión para representar regiones homogéneas y usar estructuras de datos para acelerar las operaciones de cálculo y actualización de la superficie (usando el llamado *run length encoding*, [7]).

EN PERSPECTIVA

Hemos descrito de modo bastante general cómo el problema de segmentación de regiones de interés se aborda formulando modelos de contorno activo, que confieren a cada ROI ciertas propiedades que condicionan a la solución, y visto algunos ejemplos de cómo su implementación y aplicaciones conjugan aspectos de formulación-diseño en matemáticas y computación. La idea de fondo es que si es posible formular adecuadamente un modelo de optimización con las propiedades requeridas, se puede encontrar (o acercarse) una solución siguiendo un proceso de aproximaciones sucesivas hacia el óptimo. Este proceso puede “ayudarse” con técnicas complementarias: operaciones comunes con sistemas de ecuaciones algebraicas y diferenciales requiere de métodos numéricos que garanticen convergencia y a la vez sean eficientes y tan rápidos como se pueda; el uso de técnicas de tratamiento de imágenes para resaltar características o regiones y

mejorar el ajuste de los contornos hacia zonas de borde constituye un tema por sí solo; también el aumento de resolución de las imágenes y/o contornos permite mejorar significativamente la precisión al momento de describir características de morfología como volumen, superficie o complejidad de formas; en problemas de tracking o seguimiento de objetos se emplean técnicas de segmentación que toman en cuenta movimiento y elementos como coherencia entre imágenes para aumentar el nivel de información disponible y así mejorar los resultados. En años recientes, la disponibilidad de recursos de cómputo como procesadores de múltiples núcleos o GPUs permiten realizar segmentación y seguimiento en lo que un usuario percibe como “tiempo real”, o bien automatizar y manejar eficientemente el procesamiento en escalas de terabytes o superiores, constituyendo áreas de investigación y desarrollo en la actualidad.

SOBRE EL AUTOR

Jorge Jara es alumno del Programa de Doctorado del Departamento de Ciencias de la Computación (DCC) de la Universidad de Chile, y forma parte del laboratorio de procesamiento de imágenes científicas SCIAN-Lab (www.scian.cl) a cargo del Dr. Steffen Härtel en la Facultad de Medicina. SCIAN-Lab reúne a un grupo en la interface de computación, matemáticas e investigación biomédica, y colabora estrechamente con laboratorios del CMM y el DCC formando la Advanced Imaging and Bioinformatics Initiative AIBI (www.aibi.cl).

Jorge Jara y la profesora Nancy Hitschfeld del DCC de la U. de Chile, participan como colaborador y co-investigadora en el Proyecto Fondecyt 1090246: “Partial Differential Equations for 3D Photon Denoising, Optical Flow and Adjacent Active Surface Models

for High Throughput in Vivo Spinning Disk Microscopy”, que busca desarrollar técnicas de procesamiento de imágenes para microscopía con aplicaciones biológicas.

Además forman parte del Núcleo Milenio de Morfogénesis Neuronal (NEMO, www.nemolab.cl), una iniciativa de colaboración única en Chile y América del Sur que combina experticias en neurociencia molecular, biología del desarrollo, morfogénesis, neuropatología y cuantificación de imágenes in vivo para abordar las bases genéticas de la forma, estructura y organización funcional en el sistema nervioso central. BITS

REFERENCIAS

- [1] I. Young, J. Gerbrands, and L. van Vliet. Fundamentals of Image Processing. Delft: PH publications, 1995.
- [2] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *International Journal of Computer Vision* 1:321-331, 1988.
- [3] S. Osher and J.A. Sethian. Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulation. *Journal of Computational Physics* 79: 12-49, 1988.
- [4] V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours. *Numerische Mathematik* 66: 1-31, 1993.
- [5] L. Cohen. On active contour models and balloons. *Computer Vision, Graphics and Image Processing: Image Understanding* 53(2): 211-218, 1991.
- [6] T. McInerney, D Terzopoulos. T-snakes: Topologically adaptable snakes. *Medical Image Analysis* 4(2): 73-91, 2000.
- [7] B. Houston, M.B. Nielsen, C. Batty, O. Nilsson, and K. Museth. Hierarchical RLE level set: A compact and versatile deformable surface representation. *ACM Transactions on Graphics*, 25(1):151-175, 2006.

Entrevista

Andrea Rodríguez

Por Claudio Gutiérrez

Ma. Andrea Rodríguez es profesora titular del Departamento de Ingeniería Informática y Ciencias de la Computación de la Universidad de Concepción. Tiene el título de Ingeniero Civil Informático (1989) de la Universidad de Concepción y los grados de MSc. (1997) y Ph.D (2000) en Ciencia e Ingeniería de la Información Espacial de la Universidad de Maine, Estados Unidos. Su trabajo de investigación se centra en bases de datos y sistemas de recuperación de información espacial y espacio-temporal. Ha participado o dirigido proyectos financiados por Fondecyt, ECOS/CONICY, NIMA-USA, Milenio-Mideplan y Fundación Andes. Ha publicado más de 40 artículos, escrito tres capítulos de libros y editado dos libros.



¿Cuáles son -en tu opinión- las grandes áreas básicas de nuestra disciplina. (Aquellas sin las cuales casi no se podría hablar ya de computación)? ¿Cuáles son las emergentes, aquellas a las que habría que poner más atención?

Primero habría que aclarar qué abarca la “disciplina”. ¿Estamos hablando de computación en general o hablamos de Ciencias de la Computación, Sistemas de Computación, Sistemas de Información o Tecnologías de Información?, lo que puede dar énfasis a distintas cosas.

En forma amplia, en mi opinión, las áreas básicas en computación son Algoritmos y Lenguajes de Programación, sin las cuales

no tendríamos sistemas funcionando, sistemas de manejo de información, pero por ello entiendo no sólo las bases de datos clásicas sino más bien sistemas para manejo de datos que pueden ser estructurados o semiestructurados e incluso no estructurados, y redes, lo que ha permitido la distribución y masificación del uso de sistemas computacionales.

En términos de áreas emergentes pienso que *natural interfaces*, que va desde voz, tacto, imagen, y más allá de eso, emerge y logra gran interés dentro del desarrollo impuesto por la computación ubicua. Algo que no es nuevo ni emergente, pero que sin duda debemos abordar a corto plazo es la

integración de información, entendiendo por ello la integración no sólo de distintos esquemas de un mismo tipo de datos sino distintos tipos de datos en sí (por ejemplo, texto, imágenes, voz, etc.) para lograr una infraestructura global (lo que no significa centralizada) de manejo de información.

Si tuvieras que cambiar de área por algún motivo ¿qué área elegirías y por qué?

Difícil decir, pero creo que elegiría un área que me gustara dentro de las que considere relevantes en la disciplina. Creo que igual desembocaría en Sistemas de Manejo de Información porque me gusta su combinación de aspectos teóricos con la aplicación directa en distintos dominios de información. Si pudiera comenzar de nuevo, sin embargo, buscaría de manera más temprana una preparación más formal en esta área. Creo que una sólida formación teórica permite una mayor comprensión y desarrollo posterior.

Si tuvieras que invertir en algún área de computación, ¿a qué líneas apostarías?

En tecnologías orientadas a Computación Ubicua (pervasive computing), lo que incluye muchos campos dentro de la computación, partiendo por sistemas modernos de comunicación, *natural interfaces*, e integración y servicios de información.

¿Cómo ves la Ciencia de la Computación respecto de las otras ciencias clásicas? Sus relaciones, sus conflictos, su interdisciplinariedad.

Creo que aún existe bastante desconocimiento sobre lo que esta disciplina es y abarca. Es una disciplina que lucha por ser reconocida

como una ciencia y que debe someterse muchas veces en su evaluación y valoración a estándares definidos por otras disciplinas. Sin embargo, la computación se encuentra inmersa, es usada y es relevante para muchas otras disciplinas.

La computación si bien es cierto no se asocia a necesidades básicas que sustentan la importancia de otras disciplinas como la biología, química, física, entre otras, donde el impacto es muchas veces visible en forma física y objetiva, ha alcanzado un uso tan masificado en la sociedad que forma parte de la vida cotidiana y cuyo impacto y evaluación es de alguna manera intangible y difícil de conmensurar.

¿Puedes nombrarnos -y comentarnos- algunos problemas que consideras entre los más relevantes de nuestra disciplina?

Por un lado, la masificación del uso de la computación y su rápido desarrollo abren mayores expectativas para quienes la usan y, por ende, imponen a su vez una continua demanda en aspectos tecnológicos y más teóricos. Esto puede ser visto como un problema pero también como una oportunidad de estar siempre renovando e innovando.

Por otro lado, y aunque creo que esto ha ido mejorando de alguna manera, existe mucha tendencia a ser autorreferente dentro de nuestra comunidad. Es necesario no sólo pensar que otras disciplinas necesitan de la computación sino que la computación también las necesita para entender problemas y nichos de aplicación, y desarrollo de teorías.

Por último, la computación es muy heterogénea en su metodología de investigación. Existen áreas teóricas que prueban resultados de forma analítica,

mientras otras áreas usan la evaluación experimental en base al rendimiento computacional. Existen aún otras áreas cuyo desarrollo está dirigido a dar una mayor satisfacción a usuarios finales, y que por lo tanto, se evalúan a través de técnicas menos objetivas y sistemáticas. En estos contextos se encuentran conceptos de "relevancia" y "usabilidad" que tienen un alto componente cognitivo. Esta heterogeneidad produce problemas de comunicación incluso dentro de la misma comunidad, y un problema aún mayor a la hora de comparar las distintas áreas que componen la computación.

¿Cuál piensas que es la mejor forma de desarrollar nuestra disciplina en Chile?

Por un lado, creo que debemos reforzar la formación a partir del pregrado en Ciencias de la Computación, lo que se reflejaría luego en buenos profesionales del área y serviría como sustento a una formación de posgrado e investigación. Un buen nivel de investigación es un buen signo de desarrollo, sin duda, pero por sí sola logra principalmente el reconocimiento individual de los investigadores. Yo esperaría que la disciplina se fortaleciera en forma integral de pre, posgrado, investigación e integración con el medio para proyectarse en el tiempo y no ser solamente un reflejo de las características individuales de los que forman parte de la comunidad, reducida aún, de sus investigadores.

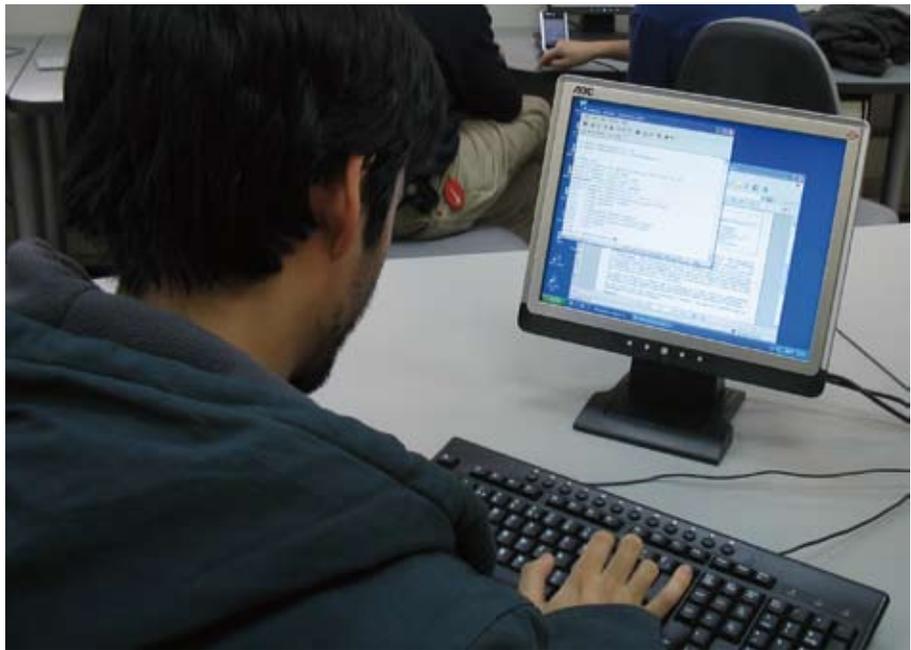
Por otro lado, creo que es importante que la academia mire un poco su entorno y logre de alguna manera un mayor impacto en el desarrollo de la comunidad o país. Creo que no sólo es importante que nosotros creamos que la investigación es importante, sino que logremos que la comunidad entienda su relevancia. Esto no es fácil, no es tampoco decir que la investigación deba estar guiada

por objetivos comerciales y de corto plazo, obviamente debe estar más allá de eso, pero sí entender que se requiere que parte de lo que se invierte en investigación tenga un retorno hacia el país.

¿Cómo evalúas el actual sistema nacional de ciencia y tecnología en relación con nuestra disciplina?

Creo el sistema nacional de ciencia y tecnología ha hecho un esfuerzo en considerar las particularidades de cada disciplina en su sistema de evaluación y distribución de recursos de investigación, pero esto aún es insuficiente. Por un lado, entiendo que el sistema busque formas objetivas de evaluación de la productividad científica de investigadores para la asignación de recursos, lo que lleva a usar ISI-Thompson como medida de la calidad de la productividad científica de investigadores frente a no muchas más alternativas. Sin embargo, esto no es suficiente y es bien claro que en nuestra área en particular existen ciertas conferencias relevantes que debieran ser consideradas tan importantes como artículos en revista. Las conferencias son necesarias para una difusión rápida de las nuevas contribuciones en nuestra área, a menos que logremos que la difusión en revistas alcance el dinamismo que otras disciplinas han alcanzado. Considerar conferencias como medio objetivo de la evaluación científica, sin embargo, requiere una evaluación seria de ellas.

Pero en general, y más allá del uso de ISI-Thompson, yo esperararía que el sistema nacional de ciencia y tecnología pudiera lograr la madurez necesaria, donde no sea necesario cada año hacer el ranking, más bien cuantitativo, de investigadores que postulan, y que un proyecto se evalúe por la calidad de la propuesta y por la capacidad de sus investigadores para llevarlo a cabo con éxito, pero siendo esto último posible de evaluar por los pares externos y un comité de especialistas en su área.



¿Qué dificultades has encontrado para desarrollarte en esta disciplina en el país?

Desde una perspectiva individual, las dificultades que he encontrado son propias de un sistema en el cual a un académico se le exigen varias cosas al mismo tiempo: docencia, investigación, ejercer tareas administrativas y en algunos casos hacer asistencia técnica. Sin claros objetivos académicos que tengan el respaldo de la institución en la cual uno se inserta, las tareas a realizar suelen ser muy dispersas y muchas veces rutinarias, lo que nos distrae de cosas más fundamentales.

Desde un punto de vista más general, existen problemas transversales a varias disciplinas que también nos afectan. Por ejemplo, los sistemas actuales de apoyo a través de becas para que buenos alumnos emigren a universidades extranjeras limitan en parte el desarrollo de nuestros programas de posgrado. Al parecer creemos que no somos tan buenos como los otros para formar a nuestros profesionales a nivel de posgrado. No quiero decir con esto que no

sea bueno otorgar oportunidades como éstas y hacer que buenos alumnos se especialicen en el extranjero en las áreas de menor desarrollo en el país, pero indudablemente hay que considerar el impacto que tiene dar un gran número de becas tanto en el fortalecimiento de nuestros programas de posgrado como en la reinserción de los que al final vuelven de estos programas. El sistema tradicional en que los alumnos consiguen financiamiento en las propias universidades extranjeras que los acogen no es algo a descartar, donde son las universidades las encargadas de “premiar” y retener a sus buenos alumnos.

¿Qué argumento darías a los jóvenes para dedicarse a nuestra disciplina?

Esta disciplina es muy amplia y por eso mismo heterogénea, donde uno puede desarrollarse en aspectos más teóricos, como son por ejemplo matemáticas discretas, hasta aspectos de gestión de información. También tiene una amplia aplicación en otras disciplinas. Una enumeración de ellas

no sería exhaustiva. Esto abre un amplio campo ocupacional a nivel profesional pero también una gran demanda de investigación. Es una disciplina donde lo abstracto más que lo tangible o físico se aplica, lo que hace que sea una muy buena alternativa para quienes deseen poner el límite de la innovación en su imaginación y creatividad. Esto es factible ya que a diferencia de muchas otras disciplinas, la computación no necesita inversiones tan grandes, comparativamente hablando, para poder lograr resultados.

¿Qué consejo darías a los jóvenes que ya cursan computación e informática para su futura carrera?

Parto diciendo que dar consejos me parece un poco presuntuoso de mi parte. Supone una cierta madurez de la disciplina que estoy tratando aún de lograr. Dicho eso, sólo puedo dar algunas ideas que no serán por cierto una lista exhaustiva de todos los consejos que pudieran recibir.

La verdad es que mis consejos son bastantes generales y no particulares a esta disciplina. Primero no pienses que dejarás de estudiar al finalizar tu carrera, más bien prepárate porque estás recién comenzando si es que quieres mantenerte actualizado. Estando aún en una situación donde tu esfuerzo tiene como mayor beneficiario a ti mismo, dirige ese esfuerzo, y no lo desperdices, úsalo para ampliar tu conocimiento y abrir tus posibilidades. En Chile existen muchos profesionales en computación, eso pone mayores exigencias para los que egresan, ya que deben buscar insertarse en un campo laboral cada vez más competitivo. Si algo he aprendido en estos años es que en este campo laboral, no sólo es necesario tener buenos conocimientos técnicos del área, tampoco son necesariamente los más brillantes los que logran desenvolverse adecuadamente, sino los que son perseverantes, responsables y comprometidos con lo que hacen.

¿Qué consejos darías a los profesores (académicos) de nuestra disciplina para desarrollar mejor su función?

Principalmente creo que debemos esforzarnos en hacer las cosas cada vez mejor. Esto es muy genérico y vago, pero si trato de ser más precisa hay que decir que siendo académicos estamos obligados a ser estudiantes de por vida, pero donde ese estudio no es para nosotros mismos sino para transmitirlo a quienes vienen atrás de nosotros en el tiempo. Más aún, es importante tratar de abrir los ojos a las futuras generaciones de manera de incentivar la exploración e iniciativa. A nivel de investigación debemos concentrarnos en la calidad de las cosas, más que en la cantidad, el problema es definir lo que es calidad y estoy segura que no será algo en lo que todos estemos de acuerdo.

¿Nos puedes contar alguna experiencia personal que pueda servir a otros colegas o estudiantes?

Una cosa que ha marcado mi desarrollo académico es el hecho que desde mi posgrado he trabajado con la comunidad de Ciencias de la Computación pero también con una comunidad multidisciplinaria dedicada a las Ciencias de Información Espacial (geográfica). Esto me ha obligado a hablar distintos lenguajes técnicos. Eso indudablemente ha enriquecido mucho mi trabajo y me ha enseñado la importancia de mirar más allá de la disciplina propia. Debo confesar que más de alguien en nuestra comunidad me preguntó en un comienzo por qué no dejaba de trabajar en cosas tan aplicadas a un dominio particular y me concentraba en el manejo más abstracto de la información. Creo que el tiempo me ha mostrado que mi opción es también válida.

Quisiera agregar que la reinserción en el país después de volver de un posgrado

en el extranjero puede ser difícil. En lo personal, no sólo regresé a Chile, sino que me inserté en un Departamento que está fuera de la capital y, que por lo tanto, tiene ciertas ventajas personales, pero también sufre de las desventajas propias del centralismo del país. Debo decir, sin embargo, que una de las cosas buenas que me sucedió fue establecer, ya sea en forma voluntaria o casual, buenos lazos con investigadores de otras universidades que realmente ayudaron a que mi inserción en la comunidad fuera mucho más fácil. No quisiera dar nombres de quienes me ayudaron, para no cometer la falta de omitir a alguien. De esto aprendí claramente que es importante tratar de conocer a quienes son tus pares y buscar puntos de interés en común. Somos una comunidad pequeña donde podemos conocernos y apoyarnos bastante. Eso debe ser aprovechado.

¿Cómo ves a la mujer en Ciencia de la Computación?

La primera experiencia que uno tiene en esta área como mujer es que generalmente se trabaja con hombres. La planta académica de nuestras universidades no ha tenido un aumento tan significativo en la participación de mujeres. Esto es aún más claro entre estudiantes, donde he visto disminuir el componente femenino. Quizás debamos acostumbrarnos al hecho de ser minorías. La participación de una mujer en computación es importante no porque sea mujer, sino por el aporte que haga a la disciplina. Esto debe colocarnos en el mismo papel que tienen los hombres. Indudablemente reconozco y he vivido las diferencias propias de nuestro género, que creo está demás enumerar y que muchas veces pueden jugar en contra de nuestra carrera académica, pero que sin duda nos complementan y fuerzan a balancear el desarrollo familiar y profesional. BITS

Entrevista

Gonzalo Navarro

Por Claudio Gutiérrez

Gonzalo Navarro es profesor titular del Departamento de Ciencias de la Computación de la Universidad de Chile, donde obtuvo el Doctorado en Ciencias mención Computación (1998), y el Magíster en Ciencias mención Computación (1995), luego de haber obtenido la Licenciatura en Informática en la Universidad Nacional de La Plata y ESLAI, Argentina, su país de origen. Es coautor del libro "Flexible Pattern Matching in Strings" publicado por Cambridge University Press. Entre 2006 y 2008 fue Director del DCC; también fue Director del Centro de Investigación de la Web (CIW), el único Núcleo Milenio en computación del país. En 2008 la Universidad de Chile lo reconoció como Mejor Docente de Pregrado y obtuvo el Premio Scopus concedido por editorial Elsevier, en virtud de sus numerosos artículos publicados y citas recibidas en su área de conocimiento. En 2010 fue elegido consejero del Consejo Superior de Desarrollo Tecnológico del Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt), cargo que ejercerá hasta el año 2013.



¿Cuáles son -en tu opinión- las grandes áreas básicas de nuestra disciplina. (Aquellas sin las cuales casi no se podría hablar ya de computación)? ¿Cuáles son las emergentes -aquellas a las que habría que poner más atención?

A mi entender, las áreas troncales en computación son Algoritmos, Lenguajes, y Software. Sin alguna de ellas difícilmente puede hablarse de nuestra disciplina. Hay muchas otras importantes, claro, pero son un poco más dependientes de la tecnología del momento.

Algunas áreas emergentes bastante evidentes son la minería de datos de todo tipo, la computación altamente distribuida y la manipulación de flujos masivos de datos. Los tres fenómenos están relacionados con el nivel de penetración que ha alcanzado la computación en nuestra sociedad.

Si tuvieras que cambiar de área por algún motivo ¿qué área elegirías y por qué?

Antes de enamorarme perdidamente y para siempre de los Algoritmos, me atrajo el tema de Bases de Datos (tanto lo formal, en el sentido de dar estructura a la información, expresividad, etc., como los problemas de eficiencia y manejo concurrente). Otro tema que siempre me gustó fue el de Compiladores y Lenguajes, pero no me veo hoy investigando en esto (me desesperan las áreas donde es difícil comparar lo que uno ha hecho con lo demás, ni teórica ni experimentalmente). Otras áreas que me atraen son la Criptografía y la Teoría de la Información. En cierto modo he logrado combinar esta última con mi área durante la última década.

Si tuvieras que invertir en algún área de computación, ¿a qué líneas apostarías?

A las áreas teóricas, porque es donde se obtienen los resultados más trascendentes en el tiempo, con lo que se puede impactar más en el prestigio académico del país. Asimismo, porque son las más exigentes intelectualmente hablando, atrayendo así a los alumnos más capaces y elevando el nivel formativo de la carrera. A la larga, esto significa tener más profesionales de muy buen nivel. Si la carrera es muy fácil, atrae a los alumnos menos capaces.

También apostaría a las áreas más aplicadas, porque es donde se puede impactar más directa e inmediatamente en el desarrollo tecnológico del país. El desarrollo de software, por ejemplo, es un área donde cualquier economía pequeña puede destacarse fácilmente, porque necesita muy poca inversión e infraestructura.

Por supuesto, hay áreas muertas que no son ni intelectualmente desafiantes, ni relevantes en la práctica.

¿Cómo ves la Ciencia de la Computación respecto de las otras ciencias clásicas? Sus relaciones, sus conflictos, su interdisciplinariedad.

Computación está en una situación muy particular. No es una ciencia clásica, del estilo de observar la naturaleza, plantear hipótesis, y verificarlas experimentalmente. Se parece más a las matemáticas, donde más bien se crean construcciones intelectuales. Pero computación es más que eso, porque tiene un componente de habilidad práctica (se puede ser un buen teórico y un mal programador) e incluso un importante componente de ingeniería (como la dirección de proyectos de desarrollo de software) que no tienen ya nada que ver con la otra punta del espectro. Como consecuencia, hay muchas actividades perfectamente válidas y de intersección vacía dentro de la disciplina: el teórico que diseña y analiza algoritmos y protocolos, el programador que implementa y experimenta, el gerente

de proyectos que dirige grandes desarrollos de software y maneja personal, metas y plazos. Ninguno de ellos tiene la menor idea de en qué consisten las otras cosas, y difícilmente entienden o aprecian lo que hacen los otros (hay excepciones, claro, de gente capaz de trabajar en un par de estas zonas simultáneamente).

Esta heterogeneidad interna, heterodoxia científica y juventud le trae muchos problemas a la disciplina. Las ciencias clásicas la miran con recelo: no se pueden tomar en serio la disciplina, salvo tal vez el área más teórica por su parecido con las matemáticas, pero esto discrimina las áreas aplicadas, que son tremendamente relevantes. Admiten su gigantesco impacto en la sociedad, pero las ven más como una ingeniería que como una ciencia, ¡y en parte lo son! No saben cómo medir la calidad de la investigación (nosotros tampoco sabemos del todo), pues las medidas clásicas no se le ajustan bien. No hay respuestas simples como factor de impacto, tasas de aceptación, etc. que sean terminantes. El solo hecho de que las conferencias representen un medio de publicación tan relevante y exigente como las revistas, o más, especialmente en las áreas más tecnológicas, es algo que en otras disciplinas no se termina de digerir. Para qué hablar de que muchas veces el producto a evaluar, nuevamente en las áreas tecnológicas, no es un artículo sino un software, un lenguaje de programación, o un protocolo (de red física, de seguridad, de Internet). No necesito decir que muchos de los logros más trascendentes de la computación son de este tipo (Internet, Web, Google... conocidos, ¿no?).

¿Puedes nombrarnos -y comentarnos- algunos problemas que consideras entre los más relevantes de nuestra disciplina?

Muchísimos, y seguramente no los conozco todos, pero puedo mencionar los relativos a las áreas emergentes que señalé antes. Uno, es que cada vez más se generan flujos de datos a escalas sin precedentes, en áreas como astronomía, bioinformática,

clima, sensores en general, etc. Se necesitan nuevos algoritmos para extraer información de esos flujos a una velocidad que muchas veces no permite ni almacenarlos. Un flujo de datos particularmente curioso es el que genera la sociedad misma con su comportamiento: compras online, páginas Web visitadas, consultas en buscadores, lo que dicen en los chats, etc. Esto se relaciona con el segundo tema, la minería de datos de todo tipo (natural y social), para extraer conocimiento a partir de ese flujo de información. El tercero tiene que ver con las limitaciones físicas a las que está llegando la miniaturización, que hace que el aumento del poder de cómputo se empiece a basar en tener muchos computadores funcionando en paralelo en vez de uno central cada vez más poderoso. Esto se da asimismo en nuevas aplicaciones donde miles de pequeños computadores (por ejemplo, sensores) trabajan en conjunto. Se necesitan nuevos algoritmos, lenguajes y protocolos para que estas aplicaciones tan complejas funcionen.

¿Cómo evalúas el actual sistema nacional de ciencia y tecnología en relación con nuestra disciplina?

Muy decente comparado con otros países de la región, pero el financiamiento es insuficiente si se compara con el de los países desarrollados, que es a lo que debemos aspirar. Los problemas relativos a la evaluación de la computación se parecen a los de otros países, aunque en algunos, donde la disciplina es más potente, han avanzado mucho más en este aspecto. En todo caso, tengo confianza en que la poca inercia propia de un sistema relativamente pequeño haga posible mejorar este aspecto en un futuro cercano.

¿Qué dificultades has encontrado para desarrollarse en esta disciplina en el país?

Partiré diciendo que Chile es uno de los contados países de la región donde se puede vivir dignamente de la investigación,

y con los recursos para llevarla a cabo exitosamente. Siento que los organismos de financiamiento, y los gobiernos en general, intentan hacer las cosas bien. Por supuesto que hay vicios, pero siento que no se supedita todo a la política; la corrupción no ha carcomido el sistema, y las peleas no son a muerte y no buscan destruir todo el sistema con tal de ganarlas. Es una tremenda diferencia con lo que se puede ver en general en el entorno.

Pero me preguntaste de las dificultades y una de las dificultades es que estamos, literalmente, en el fin del mundo. Esto es cada vez menos problema en un mundo tan interconectado, pero aún sigue siendo cierto que es más caro y pesado viajar a los centros del mundo desarrollado, o que alguien de allí viaje a Chile. No es tan común como en las universidades de Estados Unidos que algún investigador importante dicte una charla “de pasada” mientras va a otro lado. Se siente un poco este aislamiento. En este sentido, es importante que haya un financiamiento importante para viajes de los investigadores y de sus colegas, de y a Chile, más que lo usual en los países más “céntricos”.

Otra dificultad es transferir la investigación al sector privado. Lo que es común en Estados Unidos aquí es rarísimo: compañías que apuesten al desarrollo de nuevas tecnologías. Generalmente la industria local es muy conservadora, y la única forma de transferir es mediante startups, que son bastante más difíciles. Tengo la impresión de que incluso en países mucho menos estables y predecibles, como Argentina, las empresas se atreven a más. Esta es una tremenda oportunidad que Chile está dejando pasar.

Por último, la comunidad es extremadamente pequeña. A pesar de tener investigadores brillantes y una producción por cabeza muy notable, la falta de masa crítica nos juega en contra para poder postular a fondos importantes (por ejemplo, Fondap,

Instituto Milenio), y en general para desarrollar áreas de investigación a mayor escala. Se espera que crezcamos con los alumnos que formemos, además de los investigadores que traigamos del exterior. Pero los primeros no siempre encuentran fácilmente dónde insertarse. Las plazas en las mejores universidades son escasas, y en las que están surgiendo se acabarán en pocos años. Necesitamos más financiamiento para que crezcan las universidades, más centros de investigación y más compañías decididas a contratar gente de alto nivel para proyectos de verdadera innovación, si queremos que la masa crítica crezca. Si no, seguiremos el triste destino de otros países de la región, que invierten en formar excelentes investigadores y profesionales que después se irán a trabajar a los países que aprovecharán sus capacidades.

¿Qué argumento darías a los jóvenes para dedicarse a nuestra disciplina?

Si hay una disciplina que ha tenido un impacto gigantesco en la sociedad en este último siglo, es la computación. En unas décadas pasó de ser una curiosidad, a formar parte inextricable de nuestra vida. Y continuará penetrando cada vez más. Claramente esta disciplina es el futuro. Siempre tendrán trabajo, y si son buenos, habrá quien los valore.

Y a los que saben chatear, buscar en Internet, usar el Office, o enchufar la impresora, les recuerdo que eso NO es saber computación. Si creen que sí, es que no tienen la menor idea de lo que es. Esas herramientas hoy en día son simplemente la base de cualquier persona alfabetizada, y no tienen ninguna relación con las habilidades que se necesitan para “crear” esas aplicaciones. La computación es una disciplina muy desafiante intelectualmente. Tiene sus raíces en la lógica y las matemáticas (de hecho, no la recomendaría a quien no le gustara de verdad estas dos áreas), y ofrece desafíos

para todos los gustos, desde la teoría más abstracta hasta las aplicaciones que todo el mundo usa. ¡No muchas disciplinas pueden decir lo mismo!

¿Qué consejo darías a los jóvenes que ya cursan computación e informática para su futura carrera?

Que distingan entre la paja y el trigo. Que distingan los conceptos fundamentales que deben dominar toda su vida de la trivia tecnológica que ya estará obsoleta para cuando terminen su carrera. Computación es una disciplina muy dinámica y sólo lo más fundamental perdura en el tiempo. Las herramientas que verdaderamente importan son cosas como saber pensar recursivamente e inductivamente, saber generalizar y abstraer; saber modularizar, comprender el concepto de dividir y reinar, el concepto de dependencia funcional, cosas de ese nivel. Las que no importan mucho son conocer lenguajes de programación específicos, protocolos específicos, sistemas operativos específicos, herramientas de oficina y de desarrollo, etc. Sobre todo, lo más importante es que aprendan a aprender y a mantenerse actualizados. Si saben eso, aprender nuevos lenguajes, formatos, protocolos, etc. será un juego de niños.

Un segundo consejo, tan importante como el primero: aprovechen esta etapa de la vida para aprender de verdad y con profundidad. Esto no es un juego para pasar con 4.0. El que pasa con 4.0 no gana, pierde. Esta es “la” oportunidad que tienen para absorber los conceptos importantes. Nunca más volverán a tener tanto tiempo libre (aunque no lo crean hoy, tienen mucho, ¡ya verán!), tan pocas preocupaciones externas, y la mente tan fresca. Lo que aprendan ahora se les grabará para siempre. Más adelante ya no serán capaces de absorber conceptos fundamentales complejos. ¡No lo desperdicien!

¿Nos puedes contar alguna experiencia personal que pueda servir a otros colegas o estudiantes?

En 1993 había terminado mi pregrado en Argentina (en la UNLP y en la ESLAI simultáneamente, por razones demasiado largas de explicar) y llevaba un par de años trabajando en IBM. Cuando comencé a trabajar allí sentí que había tocado el cielo con las manos, pero para esta altura ya había comprendido qué rápida era la transición entre aprender algo nuevo y excitante, y pasar a dominarlo y aplicarlo monótonamente. En IBM aprendí Unix (AIX), XWindows, Motif, a jugar con el kernel con procesos, mensajes y señales (signals), y a trabajar con tarjetas de audio y video, a hacer sistemas seguros frente a fallas, y otras cosas más. Pero ya estaba claro que no aprendería nada más. Peor aún, empecé a quedarme claro que no existía una carrera de computación dentro de IBM, sino que uno debía optar por ser un programador y por ende soldado raso, o pasar a marketing, ventas y gestión. No había una carrera técnica ascendente.

Buscando dónde hacer un posgrado, que pensaba me devolvería el placer de aprender cosas nuevas, me contacté con Jorge Olivos, mi ex profesor de algoritmos de la ESLAI y hoy aún vinculado al DCC (aunque menos de lo que me gustaría). Él me propuso venir a Chile a hacer un magíster. Visité Chile en noviembre de 1993 para conocer en terreno el DCC. Jorge ya estaba algo retirado, pero me puso en contacto con un investigador llegado hacía poco de la Universidad de Waterloo y con un empuje tremendo: Ricardo Baeza Yates, que entonces se las arreglaba para ser director del DCC y hacer otras cosas al mismo tiempo. Me entusiasmó todo lo que vi, y al parecer a Ricardo también porque pronto me aceptaron como alumno.

El problema era el dinero. Yo había aparecido de la nada, así que Ricardo sólo tenía algunos restos de su proyecto Fondecyt para ofrecerme. Patricio Poblete (académico del DCC y actual director de la Escuela de Ingeniería y Ciencias de la Universidad de Chile). aportó otro poco del de él. Entre los dos, era un monto tan exiguo que no

alcanzaba para vivir (a tanta distancia creo que no es de mal gusto ilustrar los montos: recibiría alrededor de 300 mil pesos, pero un arriendo costaba 140 mil y la carrera de mi esposa, que debería recomenzarla en Chile, algo como 120 mil; ¡y aún faltaba comer, vestirse, pagar al menos la luz y el agua, etc.!). La contraparte era una cómoda situación en IBM, donde ganaba el equivalente a un millón de pesos chilenos de entonces.

En la vida hay unos pocos momentos, tal vez contados con los dedos de la mano, en que uno debe tomar una decisión verdaderamente importante, de la que depende todo el futuro. Y si la toma mal, se lamenta el resto de su vida. A veces se lamenta simplemente por no saber si la otra alternativa era mejor.

Contra toda lógica conservadora, decidí venir. Por si ese salto al vacío fuera poco, además me casé para venir los dos. Y tuve que terminar mi relación con IBM en unos términos en que, si algo salía mal, claramente no sería un lugar donde volver.

El tiempo demostraría que fue la decisión correcta, lejos. El ambiente era muy acogedor y casi familiar. Los profesores eran de primer nivel y cercanos a la vez. Con mi director descubrí lo que era la investigación y el placer de descubrir lo que nadie antes ha visto (sí, recién entonces entendí lo que venía a hacer), y hoy no concibo mi vida haciendo otra cosa. Al año, con mi Magíster terminado, fui contratado como académico (¡en esa época eso era posible!) y comencé mi Doctorado. Cada vez más cómodo en este ambiente, lo que iba a ser un Magíster de un año se convirtió en mi opción de vida. Me quedé en un lugar donde se puede vivir dignamente de la investigación.

¿Qué habría pasado de haberme acobardado? Posiblemente vegetaría en alguna oficina de IBM, lejos probablemente de ejercer la computación misma, aburriéndome soberanamente, odiándome por lo que haría para ganarme la vida, y resignado a una vida gris... y preguntándome qué habría pasado si hubiera tenido un poco más de valor.

¿Cuál es tu principal área de investigación en este momento?

He logrado congeniar dos áreas que me apasionan: las Estructuras de Datos y la Teoría de la Información. Es una interesante combinación. Las primeras son casi siempre estructuras redundantes que se crean sobre los datos puros, para facilitar su procesamiento eficiente. La segunda, el estudio de cuánta información contienen los datos, o lo que es lo mismo, cuánto es el mínimo espacio que necesito para representarlos. La belleza de esta combinación es lograr representar los datos de una manera que se acerque al mínimo que establece la teoría de la información, pero que a la vez facilite su procesamiento eficiente. Es como comprimir sus archivos y que luego de eso sea más rápido, no más lento, hacer preguntas sobre lo que contienen, manipulándolos siempre en forma comprimida en vez de descomprimirlos. En las últimas décadas se ha hecho evidente que indexar los datos (es decir, crear estructuras de datos para accederlos eficientemente) tiene mucho que ver con descubrir las regularidades que tienen, y la eliminación de esas regularidades tiene que ver con su compresión. Un eminente científico del área lo expresa como "Indexar es Comprimir".

Además del ahorro de espacio, estas representaciones comprimidas son convenientes cuando se considera la jerarquía de memoria, pues pueden caber en memorias menores y mucho más rápidas que las representaciones que necesitan más espacio. Si se logra mantener en memoria principal algo que de otro modo tendría que ser guardado en disco, habremos ganado un factor de eficiencia de hasta un millón.

Es un área de investigación verdaderamente apasionante y elegante, con teoría y con práctica combinadas, como me gusta a mí, y que recibe cada vez más atención del mundo de las aplicaciones. He estado trabajando en ella en la última década, y no parece estarse agotando ni mucho menos. BITS

Gaming for Learning: interfaces interactivas para asistir el aprendizaje y la cognición

Desde hace más de 16 años, un grupo interdisciplinario de investigadores del Centro de Computación y Comunicación para la Construcción del Conocimiento, C5, del Departamento de Ciencias de la Computación de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile, realiza investigación, desarrollo e innovación en interfaces interactivas para asistir el aprender y conocer de niños y jóvenes. En particular, el grupo Gaming for Learning de C5 ha logrado consolidar su trabajo en interacción humano-computador a través de diferentes líneas de investigación y desarrollo entre las que destacan dos de ellas: (1) Interfaces Interactivas Basadas en Sonido para Aprender y Conocer, (2) Interfaces de Videojuegos para el Aprendizaje Móvil.

El diseño y desarrollo de interfaces interactivas basadas en sonido para estimular el aprendizaje y la cognición, está centrado en niños y jóvenes ciegos. Esta línea se desarrolla desde 1994 y es pionera en su tipo, la cual considera la implementación de

una investigación cualitativa y cuantitativa sobre el aprendizaje y la cognición de aprendices ciegos a través de la interacción y utilización de interfaces de sonido 3D y háptica, generalmente utilizando videojuegos y tareas cognitivas concretas. Este trabajo implica la formulación de modelos formales de ingeniería de software para personas con discapacidad visual, el diseño, desarrollo de interfaces basadas en sonido y la evaluación de su usabilidad, para finalmente determinar el impacto del uso de estas interfaces en el desarrollo cognitivo de estas personas. Los resultados cualitativos y cuantitativos de esta investigación revelan una contribución real del uso de interfaces basadas en sonido y háptica en conjunto con las tareas cognitivas para el aprendizaje en la cognición de niños ciegos. El sonido espacial puede ayudar a mejorar y ejercitar procesos del intelecto del niño y joven ciego, tales como, la memoria háptica, memoria de corto y largo plazo, la percepción háptica, las estructuras cognitivas tempo-espaciales,



Jaime Sánchez

Profesor Titular, DCC, Universidad de Chile, Doctor en Informática y Educación, Columbia University, Nueva York; Director del Centro de Computación y Comunicación para la Construcción del Conocimiento, C5, DCC, Universidad de Chile. jsanchez@dcc.uchile.cl

la movilidad, orientación y navegación, el aprendizaje de matemáticas y de ciencia, y la resolución de problemas. Un resultado interesante y crucial de esta investigación ha sido la transferencia de este aprendizaje virtual a situaciones de la vida real. Esto nos ha llevado a confirmar la hipótesis de que los entornos virtuales basados en sonido y háptica pueden ser usados para construir significado y cognición en los aprendices ciegos.

La segunda línea de investigación, Interfaces de Videojuegos para el Aprendizaje Móvil, está centrada en niños y jóvenes videntes, y consiste en el desarrollo de modelos e interfaces de videojuegos móviles para estimular el desarrollo de habilidades de resolución de problemas y pensamiento científico en estos usuarios. Junto a ello desarrollamos y aplicamos una diversidad de métodos de usabilidad, así como también investigación de impacto del uso de estas interfaces en una diversidad de procesos cognitivos.

En ambas líneas de trabajo, el usuario final es el actor principal en el diseño, desarrollo y evaluación de las interfaces, basados en modelos y concepciones de diseño centrado en el usuario (*user-centered design*). La idea embebida en los modelos de ingeniería de software empleados es crear prototipos incrementales que son validados por los usuarios finales, minimizando errores y resultados no deseados al final del proceso, culminando así con un software entendible y usable.

También se debe tener en cuenta que para diseñar interfaces para usuarios ciegos, no basta con cerrar los ojos y ponerse en el lugar del usuario, ya que la mayoría de los elementos de interfaces y las formas de interacción que conocemos para usuarios videntes no son viables para usuarios ciegos.

Esto ocurre porque el modelo mental es diferente en cada ser humano, pero existen semejanzas entre las personas con vivencias similares y de una misma cultura. En las personas con discapacidad visual la manera de dar forma y percibir el mundo es completamente diferente a aquella de los videntes, lo que genera un modelo mental diferente. Lo anterior es sin duda la complejidad más grande que enfrentan



Grupo Gaming for Learning.

estas personas en su interacción con las nuevas tecnologías, debido a que la mayoría de sus interfaces no están diseñadas o pensadas para personas ciegas. Ello también implica un desafío no trivial para el diseño de interfaces.

EQUIPO HUMANO

El equipo humano del grupo Gaming for Learning es multidisciplinario y está compuesto por profesionales de las áreas de ingeniería civil en computación, educación, diseño gráfico y sociología.

El trabajo de ingenieros civiles en computación y estudiantes de pre y posgrado está centrado en el diseño y desarrollo de modelos formales de software, y diseño y desarrollo de las interfaces interactivas para dispositivos desktop y móviles (laptops, netbooks, pocketpcs, smartphones, iPhones) en distintos proyectos de investigación. Parte importante de la metodología de trabajo en la que participan los ingenieros es la evaluación de usabilidad de las interfaces de software en desarrollo.

Los educadores en general y educadoras especialistas en trastornos de la visión en particular, participan como apoyo en la generación de interfaces y el desarrollo de metodologías con respecto a su orientación, para fines de aprendizaje y cognición, y el diseño de instrumentos de evaluación de la usabilidad. Ellos también participan del diseño de las interfaces elaborando materiales didácticos de apoyo al trabajo con los niños y desarrollando la planificación, ejecución y evaluación del impacto cognitivo.

La construcción y aplicación de instrumentos de evaluación es realizada en conjunto con los sociólogos, quienes apoyan en el diseño y metodologías de investigación, en el análisis e interpretación de datos de las investigaciones, y en la generación de documentos y reportes. Junto con esto, los diseñadores gráficos participan en el diseño gráfico de las interfaces de software como apoyo al equipo de ingeniería, en base a los requerimientos y necesidades. En el caso de las interfaces para usuarios ciegos, realizan diseños gráficos ad-hoc para los usuarios con visión residual.

En los últimos años se han integrado neurocientistas al grupo Gaming for Learning, los cuales son especialistas en plasticidad cerebral del Medical School de la Universidad de Harvard. Su trabajo principal es estudiar científicamente los cambios y las adaptaciones que se producen en la corteza cerebral de los usuarios ciegos cuando estos interactúan con las diferentes interfaces basadas en audio y háptica, buscando establecer qué estructuras y cómo el cerebro logra adecuar y recablearse (*rewired*) para un mejor trabajo cognitivo compensando las áreas sensoriales disminuidas, basándose en modelos y teorías de plasticidad cerebral. Asimismo, se estudia cómo estos cambios a nivel de cerebro son activados por la interacción con interfaces de videojuegos y cómo pueden explicar la transferencia del aprendizaje virtual a situaciones de la vida real y sus implicaciones para fines de rehabilitación.



Usuarios jugando e interactuando con las interfaces de AudioDoom.

BACKGROUND

Desde los comienzos de esta investigación hasta la actualidad, una variedad de interfaces de software basadas en sonido y háptica han sido creadas con la finalidad de impactar el desarrollo de diferentes habilidades cognitivas. Éstas han sido evaluadas por los usuarios finales, ciegos y videntes, por medio de la aplicación iterativa de instrumentos de usabilidad para que el producto final posea una interfaz funcional, adaptada al modelo mental del usuario final y usable.

El primer software desarrollado en 1995 utilizó laberintos, personajes y objetos basados en sonido para que los niños ciegos pudieran desarrollar su memoria auditiva y habilidades espacio-temporales. Basado en el videojuego Doom de iD Software, se diseñó y desarrolló AudioDoom [1], con interfaces basadas en audio estéreo. Éste consiste en un laberinto con un pasillo principal y dos pasillos secundarios, el usuario sólo puede avanzar por laberintos encasillados. La inmersión en el ambiente se produce por medio de la ubicación izquierda, centro y derecha de los objetos y personajes dentro del laberinto. Las habilidades cognitivas estudiadas con el uso de este software fueron movilidad y orientación, estructuras espacio-temporales y navegación compleja.

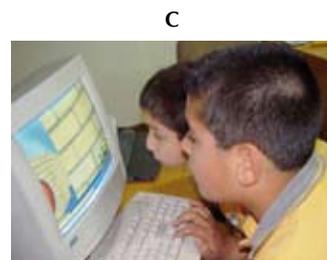
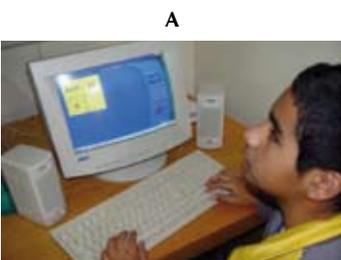
En virtud del carácter pionero e innovador de este estudio que utilizó interfaces sin pistas visuales con un extenso trabajo de usabilidad y el impacto de sus resultados, fue galardonado con los premios Stockholm Challenge Award, Suecia 2000 y Global Junior Challenge, Italia 2000.

A partir de AudioDoom como prueba de concepto y los auspiciosos resultados obtenidos, la investigación sobre interfaces basadas en sonido para el desarrollo de habilidades cognitivas fue incrementada y diversificada con variados desarrollos de ideas más complejas y extensas. El proyecto Fondecyt sobre “entornos Interactivos basados en Sonido para Aprender, TISA” (2003-2006), buscó la implementación de variadas interfaces de entornos virtuales basados en audio (estéreo, espacializado y texto) para promover el aprendizaje y el desarrollo de estructuras cognitivas. Como base para el desarrollo de aplicaciones para usuarios ciegos se plantearon dos modelos, uno de arquitectura de software y otro de desarrollo para tener una consistencia de todas las aplicaciones a desarrollar [3]. Bajo este proyecto trabajamos el aprendizaje de contenidos matemáticos con los videojuegos AudioMath, y La Granja de Theo y Seth[5]. Estos videojuegos están basados en audio estéreo y en ellos el niño aprende y practica

conocimientos de matemática básica (suma, resta, números ordinales, recta numérica). En la explicación de los conceptos básicos de formación de números se utilizan sonidos lúdicos en conjunto con verbales. Cuando el niño está en una etapa de operatoria, el sonido es principalmente verbal, ocupando el sonido lúdico como apoyo a la navegación y a la motivación por el software.

Otros procesos cognitivos estudiados fueron la resolución de problemas y las habilidades de movilidad y orientación. AudioLink [9] es un juego basado en sonido que refuerza el aprendizaje de conceptos de ciencia en un ambiente lúdico para niños ciegos, en el que interactúan con personajes y objetos con la finalidad de cumplir una misión central. Otros videojuegos basados en audio para el desarrollo de habilidades de resolución de problemas y movilidad y orientación son AudioChile y AudioVida [11].

En el marco del proyecto Fondecyt “Mundos Virtuales para la Inclusión Social, MUVIS” (2006-2008), implementamos y evaluamos la usabilidad de mundos virtuales que representan la navegación por espacios reales y que ayudan a niños y jóvenes ciegos a interactuar con su entorno real, favoreciendo así su inclusión social. El fin último del proyecto consistió en estimular el desarrollo de habilidades de resolución de problemas de la vida diaria en niños con discapacidad visual, por medio del uso de interfaces sin pistas visuales basadas en audio para que puedan resolver problemas reales en contextos reales, desarrollando habilidades tales como: colaboración, navegación, movilidad y orientación. Se elaboró un modelo de desarrollo de aplicaciones de interfaces móviles por medio de una descripción teórica/abstracta del mismo: el paradigma de especificación, la funcionalidad, la documentación



Usuarios ciegos interactuando con (A) AudioMath, (B) La Granja de Theo y Seth, (C) AudioVida y (D) AudioChile.

conceptual, y la arquitectura estructural de los componentes, sus restricciones y extensiones [13]. Aplicaciones de interfaces móviles como mBN [4], ambientGPS [7], AudioTransantiago [4], EMO [8] y PYOM [2] fueron generadas para el apoyo y asistencia in situ de jóvenes ciegos en su navegación por el metro, el barrio, la locomoción colectiva, y en edificios como su escuela, respectivamente.

El proyecto “Inclusión Digital para Aprender Ciencias, IDAC” [12], con financiamiento del Ministerio de Educación, consistió en diseñar e implementar un videojuego de rol educativo y colaborativo, para ser utilizado en computadores personales. Con este videojuego se pretendió integrar a usuarios videntes con usuarios ciegos en situaciones de aprendizaje, poner en práctica sus conocimientos en ciencias y desarrollar habilidades de resolución de problemas.

Por otra parte, durante los años 2007 y 2008 se desarrolló el proyecto “Aprendizaje de la Biología con Tecnología Móvil, ABTm” [14] (financiado inicialmente por Microsoft y luego por el Ministerio de Educación), cuyo objetivo fue desarrollar videojuegos en dispositivos móviles para el desarrollo de habilidades de resolución de problemas en ciencias con aprendices de octavo año básico. Estos videojuegos móviles aprovechan el contexto para generar espacios de aprendizaje fuera de la sala de clases, como un zoológico o un museo. La metodología consistió en actividades preparatorias con el profesor, diseño de actividades de trabajo en el aula, actividades en terreno con videojuegos de trivia para pocketPC y una actividad central con un videojuego móvil de estrategia (“Evolución”), también para pocketPC y classmate. Ambos videojuegos aprovechan las cualidades touch de la interfaz de la pocketPC, en la que el alumno



Usuarios ciegos interactuando con las interfaces móviles desarrolladas.

utiliza el stylus para interactuar y avanzar en las diferentes etapas. El videojuego presenta una interfaz atractiva, lúdica e intuitiva, siendo un elemento clave para la experiencia del juego, ya que es de vital importancia retener la atención del usuario. Para esto se reutilizaron conceptos gráficos y de interacción de este tipo de juego, favoreciendo la comprensión de la interfaz.

ACTUALIDAD

En la actualidad se está ejecutando el proyecto Fondecyt “Juegos basados en Audio y Háptica para la Movilidad y Orientación, JAHMO” (2009-2011). En el primer año de proyecto elaboramos un modelo de ingeniería de software para el desarrollo de aplicaciones móviles basadas en videojuegos, de manera de mejorar habilidades de movilidad y orientación en

usuarios ciegos [6]. Bajo este modelo, se han construido dos videojuegos sin pistas visuales basados en tecnología de audio y háptica para estimular el desarrollo de habilidades de movilidad y orientación en los niños ciegos. En la actualidad se están realizando evaluaciones cognitivas para conocer la efectividad e impacto del uso de estas interfaces de videojuegos en el desarrollo de las habilidades cognitivas estudiadas.

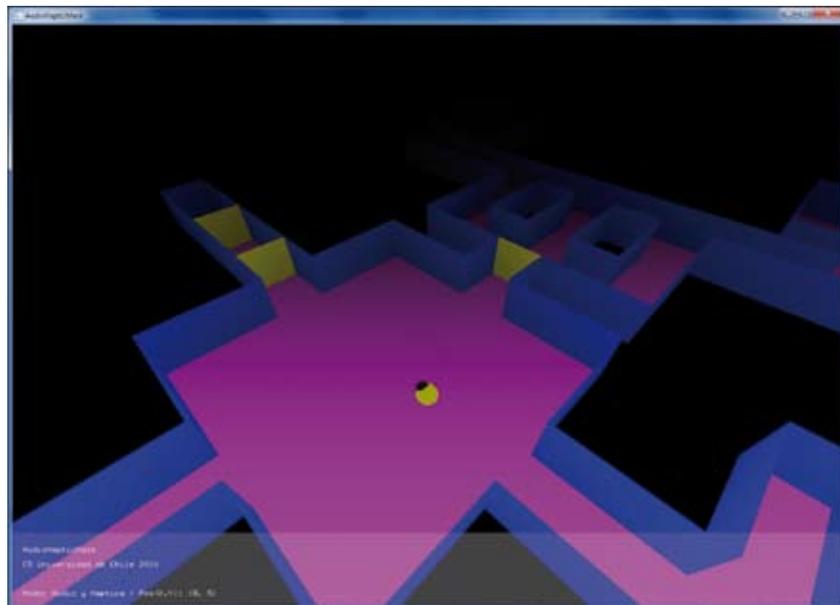
Al mismo tiempo, en conjunto con investigadores de la Pontificia Universidad Católica de Valparaíso estamos desarrollando el proyecto “Navegación para Aprendices Ciegos a través de Videojuegos, NACVI”, con el propósito de evaluar e investigar el impacto del uso de videojuegos basados en audio, en el desarrollo y uso de habilidades de Movilidad y Orientación en situaciones de espacios cerrados desconocidos en niños en edad escolar.



Usuarios jugando en el BuinZoo y en clases con “Evolución”.

En la actualidad desarrollamos el proyecto Fondef TIC-EDU “Videojuegos para el Desarrollo de Habilidades en Ciencia a través de Celulares, ViDHaC2”. Este proyecto busca abordar los problemas de los bajos resultados en el aprendizaje de la ciencia entre los alumnos y las debilidades en didáctica entre los profesores de ciencia. El profesor utiliza una interfaz Web para diseñar y desarrollar un videojuego tipo Role Playing Game (RPG) que luego es descargado y jugado por sus alumnos en el celular. En el videojuego los alumnos interactúan por medio de un personaje en un mundo virtual realizando tareas, respondiendo preguntas e interactuando con otros personajes para finalmente avanzar y ganar el juego. De esta forma, se busca aprovechar las oportunidades que ofrece la masividad de los teléfonos celulares, que permite una relación alumno/dispositivo cercana al 1:1, la posibilidad de disminución de brecha digital y el aprovechamiento de la movilidad propia de estos dispositivos. El conjunto de estos elementos contribuirá a un cambio metodológico profundo en el aprendizaje de la ciencia, más cercano a los estilos de aprendizaje de la actual generación de aprendices. Además, el proyecto contribuye al fortalecimiento de la capacidad tecnológica del país desarrollando el mobile learning o mLearning y con potencial de escalamiento.

Finalmente, y en colaboración con los Profesores Álvaro Pascual-Leone y Lotfi Merabet del Departamento de Neurología y Oftalmología del Harvard Medical School, se está ejecutando el proyecto “Audio-Based Navigation in the Blind” que es financiado por el National Institute of Health (NIH), por un plazo de cinco años a contar de 2009. El trabajo en conjunto tiene por finalidad diseñar y desarrollar interfaces de videojuegos basadas en sonido para estimular habilidades de navegación complejas (movilidad y orientación), para luego estudiar su impacto en la plasticidad cerebral del joven ciego y los cambios adaptativos de las diferentes áreas de la corteza cerebral, cuando el usuario juega e interactúa con el videojuego, todo esto determinado mediante la utilización de imagen por resonancia magnética funcional (fMRI). ^{BITS}



Captura de pantalla de AudioHapticMaze, videojuego desarrollado en el proyecto JAHMO.

REFERENCIAS

- [1] Lumbreras, M. & Sánchez, J. (1999). Interactive 3D sound hyperstories for blind children. Proceedings of the ACM-CHI '99, Pittsburgh, USA, pp. 318-325.
- [2] Sáenz, M., Sánchez, J. (2010). Indoor Orientation and Mobility for Learners Who are Blind. In Brenda K. Wiederhold, Giuseppe Riva, Sun I. Kim (Editors), Annual Review of Cybertherapy and Telemedicine 2010, Imaging the Future, Chapter 32, summer 2010. San Diego, CA: Interactive Media Institute, Volume 8, pp. 131-134.
- [3] Baloian, N., Luther, W. & Sánchez, J. (2002). Modeling educational software for people with disabilities: Theory and Practice. Proceedings of the Fifth International ACM SIGCAPH Conference on Assistive Technologies, ASSETS 2002, pp. 111-118, Edinburg, July 8-10, 2002.
- [4] Sánchez, J. (2007). Sound & learning in blind children: A case of a real contribution of ICT to learning. Proceedings of the CAL'07 conference, 'Development, Disruption and Debate'. Ireland, 26- 28 March, 2007, pp. 070.
- [5] Sánchez, J. (2008). User-Centered Technologies for Blind Children. Human Technology Journal, 45(2), November 2008, pp. 96-122.
- [6] Sánchez, J. (2010). A Model to Develop Videogames for Orientation and Mobility. 12th International Conference on Computers Helping People with Special Needs, ICCHP 2010, Vienna, Austria, July 12-13, 2010.
- [7] Sánchez, J., De la Torre, N. (2010). Autonomous Navigation through the City for the Blind. 12th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2010, Orlando, Florida, USA, October 25 - 27, 2010, pp. 195-202.
- [8] Sánchez, J., Elías, M. (2007). Guidelines for Designing Mobility and Orientation Software for Blind Children. In C. Baranauskas et al. (Ed.): INTERACT 2007, Lecture Notes in Computer Science LNCS 4662, Part I, pp. 375-388, 2007. © IFIP International Federation for Information Processing.
- [9] Sánchez, J., Elías, M. (2007). Science Learning by Blind Children through Audio-Based Interactive Software. Annual Review of CyberTherapy and Telemedicine: Transforming Healthcare through Technology, Volume 5, pp. 184-190.
- [10] Sánchez, J., Mendoza, C., Salinas, A. (2009) Mobile serious games for collaborative problem solving. Annual International CyberTherapy and CyberPsychology Conference 2009. Lago Maggiore, Verbania_Intra, Italy, 21-23 June 2009.
- [11] Sánchez, J., Sáenz, M. (2006). 3D sound interactive environments for blind children problem solving skills. Behaviour & Information Technology, Vol. 25, No. 4, July – August 2006, pp. 367 – 378.
- [12] Sánchez, J., Sáenz, M. (2009). Video Gaming for Blind Learners School Integration in Science Classes. In T. Gross et al. (Eds.): INTERACT 2009, Part I, LNCS 5726, pp. 36–49, 2009. © IFIP International Federation for Information Processing 2009.
- [13] Sánchez, J., Sáenz, M., Baloian, N. (2007). Mobile Application Model for the Blind. In C. Stephanidis (Ed.): Universal Access in HCI, Part I, HCII 2007, LNCS 4554, pp. 527–536, 2007. © Springer-Verlag Berlin Heidelberg 2007.
- [14] Sánchez, J., Salinas, A., & Sáenz, M. (2007). Mobile Game-Based Methodology for Science Learning. In J. Jacko (Ed.): Human-Computer Interaction, Part IV, HCII 2007, LNCS 4553, pp. 322–331, 2007.

CONFERENCIAS:

Alberto Mendelzon Workshop en Fundamentos de Bases de Datos (AMW 2011)

Santiago, Chile, 9-12 Mayo 2011

<http://db.ing.puc.cl/amw2011/>

Este Workshop es una iniciativa de la comunidad latinoamericana de investigadores identificados con las áreas de bases de datos, manejo de información y la Web, a la cual nuestro amigo Alberto Mendelzon contribuyó con enorme generosidad. Nuestro objetivo es establecer en el cono sur una instancia científica periódica de alto nivel en los aspectos fundamentales del área. Creemos que ésta es una excelente forma de mantener viva la memoria de Alberto, y a la vez de incrementar y solidificar la investigación científica en la región. Como en los dos eventos anteriores, otra de nuestras principales motivaciones es ayudar a que los alumnos latinoamericanos interesados en los fundamentos de bases de datos y la Web (especialmente los alumnos de posgrado) tengan la oportunidad de interactuar con algunos de los más destacados expertos mundiales del área.

Este año el Alberto Mendelzon Workshop se realizará en Santiago, Chile, en el Centro de Extensión de la Pontificia Universidad Católica de Chile, desde 9 al 12 de mayo de 2011. El Workshop consistirá en cuatro charlas invitadas más los artículos aceptados por el comité de selección. Además, los días 14 y 15 de mayo, se realizará un Workshop de estudiantes graduados.

Este año tenemos un destacadísimo grupo de charlista invitados, que se compone de los profesores Marcelo Arenas (Pontificia Universidad Católica, Chile), Stefan Decker (DERI Institute), Héctor García-Molina (Stanford University) y Georg Gottlob (University of Oxford).

Además contamos con el siguiente comité de programa que reúne a destacados expertos del área:

Chairs:

Pablo Barceló (Universidad de Chile)
Val Tannen (University of Pennsylvania)

Comité de programa:

Virgilio Almeida (UFMG, Brazil)
Renzo Angles (Universidad de Talca, Chile)
Paolo Atzeni (Universidad Roma Tre)
Peter Buneman (University of Edinburgh)
Benjamin Bustos (Universidad de Chile)
Diego Calvanese (Free University of Bozen-Bolzano)
Carlos Castillo (Yahoo! Research)
Nilesh Dalvi (Yahoo! Research)
Susan Davidson (University of Pennsylvania)
Amol Deshpande (University of Maryland)
Alin Deutsch (UC San Diego)
Luna Dong (AT&T Labs)
Wenfei Fan (University of Edinburgh)
Todd J. Green (UC Davis)
Mauricio Hernández (IBM Almaden)
Ihab F. Ilyas (University of Waterloo)
Marta Mattoso (UFRJ, Brazil)
Tova Milo (Tel Aviv University)
Filip Murlak (University of Warsaw)
Gonzalo Navarro (Universidad de Chile)
Flavio Rizzolo (Carleton University)
Yasin N. Silva (Arizona St. University)
Caetano Traina Jr. (USP, Brazil)
Jan Van den Bussche (Universiteit Hasselt)

Victor Vianu (UC San Diego)

María-Esther Vidal (Universidad Simón Bolívar, Venezuela)

Jef Wijsen (Université of Mons)

Organización local:

Claudio Gutiérrez (Universidad de Chile)
Jorge Pérez (Pontificia Universidad Católica, Chile)

Fechas importantes:

Deadline: 22 de Febrero, 2011

Notificación: 31 de Marzo, 2011

Buscamos artículos originales en los aspectos fundamentales de bases de datos, incluyendo las siguientes (entre otras):

Computational model theory; constraint databases; data exchange; data integration; data mining; data modeling; data management on the Web; data streams; data warehouses; database algorithms; distributed and parallel databases; incompleteness; inconsistency; information retrieval; integrity constraints; knowledge bases; logic; multimedia; physical design; privacy; provenance; quantitative approaches; query languages; query optimization; ontologies; real-time data; recovery; scientific data; search; security; semi-structured data; sensor data; services; spatial data; temporal data; transactions; uncertainty; updates; views; workflows; XML.

Los artículos deben estar escritos en Latex, en estilo LNCS, con una extensión máxima de doce páginas. Para enviar artículos utilizar EasyChair. Los artículos serán publicados en formato electrónico en CEUR Proceedings.

Pablo Barceló
DCC, Universidad de Chile

VEN A CONSTRUIR CON NOSOTROS EL FUTURO.

EXPLORA LA FRONTERA DE LAS CIENCIAS DE LA COMPUTACIÓN E INFORMACIÓN.

PROGRAMAS DE POSGRADO

DCC UNIVERSIDAD DE CHILE:

- Doctorado en Ciencias mención Computación
- Magíster en Ciencias mención Computación
- Magíster en Tecnologías de la Información





MÁS INFORMACIÓN: <http://www.dcc.uchile.cl/posgrado>



Departamento de Ciencias de la Computación
UNIVERSIDAD DE CHILE

REVISTA
BITS de Ciencia
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

UNIVERSIDAD DE CHILE



fcfm

Ciencias de la
Computación
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

www.dcc.uchile.cl/revista

revista@dcc.uchile.cl