

# DOCTORADOS DEL DCC



## HÉCTOR FERRADA



Mi trabajo de Doctorado se enmarca en el campo de Document Retrieval (DR), el cuál se enfoca en el diseño de representaciones de conjuntos de textos (documentos), que usualmente son un gran número, a fin de responder en tiempo real y rápidamente a consultas por documentos. En esta área el principal objetivo es que las soluciones puedan llevarse a cabo sobre estructuras que respondan más rápido y/o requieran de menos espacio en memoria principal que lo obtenido por las soluciones existentes, considerando para ello tanto la teoría como la práctica.

Mi trabajo de tesis considera al menos con la misma relevancia que el dado por DR al tiempo de respuesta de la consulta, que el espacio en memoria principal de la estructura final resulte ser tan pequeño como su cota inferior lo permita, y que idealmente su construcción también sea eficiente en los mismos términos de espacio y tiempo. Dentro de este marco general, abordamos el caso cuando los documentos provienen de una gran colección de textos arbitrarios. Este escenario ocurre con colecciones de secuencias biológicas y repositorios de software, entre otros. Existen varias estructuras de datos comprimidas para DR para hacer frente a este problema. Sin embargo, en la práctica las propuestas con el mejor rendimiento en tiempo requieren de demasiado espacio adicional con respecto al mínimo teórico conocido.

Las principales contribuciones de mi tesis fueron las siguientes: (1) construimos índices para DR en base a la compresión Lempel-Ziv 1978 en lugar del clásico arreglo de sufijos; (2) manipulamos colecciones altamente repetitivas en base a la compresión Lempel-Ziv 1977; (3) iniciamos un estudio para entregar respuestas aproximadas a problemas de DR. Como

principales resultados obtenidos destacan nuestros índices para DR basados en la compresión Lempel-Ziv, siendo las primeras propuestas en incorporar estos esquemas de compresión al campo de DR. Nuestras estructuras resuelven los dos problemas fundamentales en este campo: Document Listing (DL) y Top-k Document Retrieval (Top-k). Por otro lado, ofrecemos estructuras que entregan respuestas aproximadas a dichos problemas, ahorrando considerable espacio y/o tiempo comparado con cualquier otra estructura que entregue una respuesta completa a alguno de estos problemas. Finalmente, ofrecemos también contribuciones en indexación de texto y la mejor implementación existente para consultas de valores mínimos en rangos (Range Minimum Queries).

Los inicios de mi estadía en el Doctorado no fueron fáciles. En primer lugar, tuve que lidiar con el problema de descubrir mi área de investigación. Me involucré en diversos cursos para ver en qué podría trabajar, hasta que finalmente comencé a trabajar con el Profesor Dr. Gonzalo Navarro en el área de Algoritmos. Además, siempre tuve muchos problemas con el inglés, lo cual me paso la cuenta en el DCC. También venía de trabajar ya más de cuatro años en la industria, lo que creo que me sacó del ritmo de estudios que tenía en pregrado. Otro gran desafío fue el desarrollar habilidades para investigar; por ejemplo, cómo desarrollar bien una idea novedosa. Es decir, primeramente que ésta madure lo suficiente (y en poco tiempo) como para redactarla, defenderla y presentarla a fin de convencer a otros que ésta vale la pena y así continuar con su desarrollo y posible implementación. Además se requiere de gran capacidad, estudios y práctica para lograr redactar adecuadamente una nota científica que seduzca al lector.

Actualmente me encuentro realizando un Postdoctorado en Finlandia, en la Universidad de Helsinki, en el grupo de Genome-Scale Algorithmics. Estaré en ese trabajo hasta fines del 2017 y después de eso quisiera ejercer en mi país. ■



IMAGEN 1. HÉCTOR FERRADA (AL CENTRO) JUNTO A SU COMISIÓN EVALUADORA. DE IZQUIERDA A DERECHA, PROFESORES: CLAUDIO GUTIÉRREZ, DIEGO ARROYUELO, GONZALO NAVARRO Y BENJAMÍN BUSTOS.

## CRISTÓBAL NAVARRO



Soy Ingeniero Civil en Informática de la Universidad Austral de Chile y recientemente Doctor en Ciencias de la Computación del Departamento de Ciencias de la Computación (DCC) de la Universidad de Chile. Mi supervisor de Doctorado en el DCC fue la Profesora Nancy Hitschfeld del DCC, con quien trabajé en el problema de cómo acelerar algoritmos geométricos usando procesamiento por GPU. La experiencia de trabajar con la profesora Nancy fue totalmente gratificante, lo cual me dejó motivado para seguir el camino de la investigación. Durante 2011 comencé además a trabajar en temas de investigación interdisciplinaria de física computacional con el Dr. Fabrizio Canfora.

Me gustaría decir que escogí el DCC para realizar mi Doctorado no solo por su excelente nivel orientado a la investigación en Ciencias de la Computación, sino también por la calidad humana de los académicos, secretarías y técnicos. Es sabido que ir a otra ciudad a hacer un postgrado es un desafío en varios aspectos, sin embargo la Universidad y los amigos que hice me hicieron sentir como en casa. Mi estadía en el DCC fue un período de crecimiento académico y personal que requirió de mucha dedicación y una actitud perseverante. Pude formar fuertes lazos de amistad con compañeros de distintos países, así como conocer nuevos colegas, lugares y culturas de otros países.

Mi tesis de Doctorado corresponde a una investigación interdisciplinaria que combina el área de algoritmos paralelos con modelos de *spin* de la física teórica. Un modelo de spin es un modelo matemático que comenzó siendo usado para entender el comportamiento magnético y termodinámico de un sistema de partículas (spins) fijas que interactúan energética-

mente entre ellas en base a una definición microscópica del sistema. Todo esto queda bien representado por un grafo, donde los vértices corresponden a estados y los arcos representan las interacciones. Hoy en día, se sabe que los modelos de spin no son importantes solo en física, sino también pueden modelar problemas de optimización, crecimiento celular o incluso comportamiento de redes sociales. En particular, dado un grafo donde las interacciones pueden ser definidas por un Hamiltoniano, es posible obtener características a nivel general que no son triviales de predecir.

Mi tesis propone dos algoritmos paralelos para acelerar el estudio de modelos de spin. El primero es un algoritmo paralelo exacto que aprovecha ciertas simetrías que emergen en el espacio de configuraciones, pudiendo así procesar múltiples elementos por el costo de uno, y al mismo tiempo realizar este proceso en paralelo. El segundo algoritmo propuesto es un algoritmo Monte Carlo paralelo que acelera la convergencia de los cálculos utilizando un método adaptativo. La principal ventaja del algoritmo es que ajusta de forma dinámica los parámetros sobre cada grafo para así poder realizar una simulación más eficiente en paralelo. Ambos resultados brindan mejoras notorias en rendimiento, abriendo la posibilidad de poder estudiar grafos más complejos y de mayor tamaño.

Desde marzo de 2016 soy académico de la Universidad Austral de Chile, en Valdivia. La mitad del tiempo avanzo en un proyecto de postdoctorado Fondecyt y la otra mitad me dedico a hacer docencia en cursos relacionados con algoritmos paralelos y computación gráfica. Mi plan académico a futuro es quedarme en Valdivia, seguir avanzando en mi postdoctorado, y realizar estadías en el extranjero para formar nuevos vínculos de investigación. En el futuro trataré de transmitir a mis estudiantes el gusto por la investigación, ayudando así a impulsar la Ciencia de la Computación en el sur de Chile. ■



IMAGEN 2.  
CRISTÓBAL NAVARRO (AL CENTRO) JUNTO A SU COMISIÓN EVALUADORA. DE IZQUIERDA A DERECHA, PROFESORES: ROMAIN ROBBES, FABRIZIO CANFORA, NANCY HITSCHFELD, GONZALO NAVARRO Y RODRIGO SOTO.

## MIGUEL ROMERO



Finalicé mi Doctorado en el Departamento de Ciencias de la Computación (DCC) de la Universidad de Chile, en agosto de 2016 bajo la supervisión del Profesor Pablo Barceló. Desde el comienzo recibí el apoyo de mi supervisor en todo sentido. Cada vez que quería discutir alguna idea, podía ir a su oficina y él se daba el tiempo para conversar. Además me ayudó a hacer contactos con investigadores extranjeros,

lo cual me permitió realizar varias pasantías de investigación. Esto fue fundamental para el desarrollo de mi tesis, ya que ésta se basó en varias ideas que nacieron de estas colaboraciones. También me gustaría destacar la gran calidad humana de mis compañeros, de todo el cuerpo docente y del cuerpo administrativo del DCC, los cuales hicieron que toda esta experiencia fuera realmente excelente.

Mi tesis se enmarcó en el área de bases de datos de grafos, la cual ha ganado mucha relevancia hoy en día debido a aplicaciones como la Web Semántica o las redes sociales. Nuestro foco fue estudiar la complejidad de dos lenguajes fundamentales para grafos: las *Unions of Conjunctive two-*

*way Regular Path Queries* (UC2RPQs) y las *Regular Queries* (RQs), las cuales han sido propuestas recientemente como una extensión natural de las UC2RPQs.

Nuestra principal contribución fue obtener nuevos algoritmos para el problema de evaluación y contención de estos lenguajes. En particular, definimos restricciones de UC2RPQs que se pueden evaluar más eficientemente que el caso general (evaluar UC2RPQs es NP-completo en general, y luego computacionalmente intratable). También demostramos que la complejidad de contención (o implicancia lógica) de RQs no es significativamente más alta que la del caso de las UC2RPQs. Esto último muestra que las RQs alcanzan un buen balance entre expresividad y complejidad.

Después de mi Doctorado, en el segundo semestre de 2016, realicé una estadía de investigación en el Simons Institute for the Theory of Computing, en la Universidad de Berkeley. Estuve asociado al programa “Logical Structures in Computation”. Esta estadía fue muy positiva para ampliar mis conocimientos en lógica computacional y para conectarme con investigadores de todo el mundo. Actualmente estoy realizando un postdoctorado en la Universidad de Oxford bajo la supervisión de Standa Zivny en el área de satisfacción de restricciones. ■



IMAGEN 3.

MIGUEL ROMERO (AL CENTRO) JUNTO A SU COMISIÓN EVALUADORA. DE IZQUIERDA A DERECHA, PROFESORES: GONZALO NAVARRO, JORGE PÉREZ, PABLO BARCELÓ Y AIDAN HOGAN.