

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 interdisciplinaria?

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, empezó 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