

# Alberto visita Chile

## (una visión muy personal)

claudio gutierrez / Junio 2005

No aparentaba en el trato diario el gran hombre que era. El esforzado vendedor callejero de sopaipillas de Blanco Encalada con Beaucheff, acostumbrado a atender bajo la lluvia a su clientela de obreros, mecánicos y empleadas domésticas, nunca imaginó que su cliente favorito ese invierno era uno de los científicos más grandes del mundo que llegaría a ser elegido *Fellow* de la *Royal Society of Canada*. Supongo que de haberlo sabido, le habría cambiado el aceite a su paila y preparado sus exquisiteces con menos grasa y más zapallo. Su cliente bajaba regularmente del edificio del Departamento de Computación a pedirle unas “tortas fritas”. Dedujo que era extranjero pues el primer día casi le vació el frasco de ají para untar una sopaipilla. “Estaba bastante picante” nos confesaría más tarde Alberto restándole importancia al incidente.

Las semanas que estuvo en Chile, Alberto se alojó en un hotel de Providencia. Venía regularmente al Departamento. En un momento de euforia, recuerdo, imaginamos un cursillo de bases de datos dictado por él, y administramos con voluntarismo esas sanas ilusiones. Así tuvimos el lujo de tener a uno de los maestros mundiales del tema dando un curso para dos alumnos: una que la habían forzado a ir, y otro que aún hoy que cada vez que me encuentra en el pasillo, me repite avergonzado: “¡Por qué no me dijiste quien era Mendelzon!”. Yo comencé a ir con el objetivo de hacer número, con esa displicencia propia de los ignorantes, hasta que el maestro me cautivó con sus enfoques y motivaciones. El curso aquel me recordó una anécdota que contaba un tanguero uruguayo sobre su “compatriota” Gardel, cuento que viniendo de uruguayos es probable que sea ficción. Alguna vez el Zorzal criollo daba un concierto en el

Colón en Buenos Aires, y por alguna desinteligencia en el horario, llegó sólo un matrimonio de ancianos. Por supuesto los *managers* y administradores ya estaban cancelando la función, cuando Gardel apareció en el escenario, y comenzó a cantar. Dió esa noche, dicen, uno de sus conciertos más largos y emotivos, como si el Colón estuviese repleto. Nosotros tuvimos el privilegio de tener nuestro Gardel de las Bases de Datos, desplegando cuatro semanas su profesionalismo en una pequeña aula.

Entretanto, con Carlos Hurtado, su ex-alumno de OLAP en Toronto, *Stolen Charles* para el traductor de Google y Carlitos para nosotros, nos preparábamos frenéticamente para aprovechar al maestro. Fuimos al tercer día su oficina a presentarle algunas ideas vagas sobre un modelo de datos semántico que ebullía en las discusiones del Consorcio de la Web. Alberto, por supuesto, no sólo ya había escuchado de RDF, sino que dominaba el tema. Sus preguntas fueron al hueso, sorprendiéndonos a nosotros que pensábamos eramos los expertos y esperábamos una discusión general. Y por si no eso bastara, nos pasó varios papers ya impresos de antemano y una decena de punteros a otros. Salimos desmoralizados, aunque con una brasita encendida: el tema no era inútil, y Alberto delicadamente se había preocupado de mostrar suficiente interés en el tema. Conocía perfectamente el límite al que se puede llevar la crítica descarnada para que haga su efecto y a la vez evite que los involucrados se depriman. Nos tomó un par de semanas recuperarnos y darnos ánimo para volver a la carga. Esta vez resultó todo un éxito. Alberto se entusiasmó con la idea. ¡Pero muy tarde! Por esos días ya volvía a Canadá. La interacción siguió por *email* y la idea cristalizó en lo que él mismo terminó bautizan-



Figura 1: Alberto en el lago Colbún

do como *Foundations of Semantic Web Databases*.

Pero volvamos a las andanzas de Alberto por Chile. El Centro de Investigación de la Web, que por iniciativa de Ricardo Baeza lo había traído a Chile, organizó un asado campestre en un fundo en Talca (que Alberto siempre llamó “estancia”...). Por esos azares de la vida, yo quedé a cargo de llevarlo en mi auto. Allí se encuentra con Cecilia, que tenía por esos días 7 meses de embarazo, Gonzalito, de cuatro años, y Pedro que recién comenzaba a balbucear, datos que sumados a las dimensiones del auto dan una idea de su paciencia de santo. Contra todos los pronósticos, fue un viaje ideal: todos los pasajeros contentos, Cecilia aclarándole algunas dudas a un muy bien informado turista, quien alternaba con los niños, y re-

spondía infatigablemente al “Alberto, ¿qué es esto?” seguido por una secuencia no siempre finita de “¿por qué?”

Llegamos a Talca, cuya asociación con París y Londres –que Carlos tuvo el coraje de documentar en su tesis– le simpatizó bastante a Alberto, y enfilamos a la cordillera, para llegar a la “estancia”. Nos esperaban allí Carlos y Gonzalo Navarro con sus respectivas esposas, y... más niños. Alberto paseó por la laguna, inutilizó un telescopio debido al exceso de celo investigativo, y comió un asado de cabrito con un apetito que los parroquianos no veían hace décadas. También fuimos a la cordillera, casi al límite con Argentina, donde cada afirmación imprecisa sobre la geografía local de Carlos o mía era refutada por Alberto que



Figura 2: De izquierda a derecha: Carlos, Alberto, Carlos Diaz, Claudio, Gonzalo

ya había leído la versión correcta en su *tourist guide*. Preguntó por los detalles de cada comida que tuvo delante (aprendió que un modesto postre de leche puede presentarse como *crème brûlée*), anduvo a caballo y pidió abreviar el paseo por un dolor de espalda (la única manifestación de displacer que le vi en toda su estadía), y comió, comió, comió, como sólo lo puede hacer alguien que está en paz con su espíritu. Recuerdo que al irse a dormir –alojamos allá esa noche– prefirió llevarse unos viejos *New York Review of Books* que encontró en vez de los magníficos catálogos de caballos chilenos que le ofrecimos.

Se trajo de recuerdo un estribo de madera típico de la zona que le regaló un criador de caballos, unas botellas de la cosecha del dueño del fundo (y del ex-

telescopio...), y salimos temprano a visitar las viñas de la zona, donde catamos todos los mostos que se nos cruzaron por delante... De vuelta a Santiago, se cantó todo el viaje a dúo con Julio Sosa, incluyendo los gestos técnicos de tanguero profesional. Todavía me parece escuchar su notable interpretación de *Tarde*, con un “...de cada amor que tuve tengo heridas, heridas que no cierran y sangran todavía...”, sólo superado por mi suegro inspirado por el whisky y los ventosos y románticos atardeceres montevideanos.

En el camino, pasamos a un curioso museo en Santa Cruz, “el mejor museo de Chile” según sus organizadores (“y además un buen Museo” agregó Alberto) donde reímos mucho viendo a los ancestros de Carlos detrás de las vidieras. “Y pensar que Carlos tenía esta



Figura 3: Los cuatro jinetes... de izquierda a derecha: Claudio, Carlos, Gonzalo y Alberto

familia en Chile” decía, como un descubrimiento curioso y simpático a la vez. Ya de noche, recuerdo que nos pidió encender la radio para ir escuchando los resultados parciales de las elecciones presidenciales argentinas que se desarrollaron ese domingo.

Cuando llegamos a Santiago de vuelta, comentamos con Cecilia –que aún estaba impresionada por su sencillez– la mala ocasión en que nos visitaba, pues no podíamos despegarnos de los niños para poder sacarlo a pasear por Chile como se merecía.

Sin embargo, lo invitamos a salir un par de fin de semanas y feriados más. Fuimos a un circuito por la playa, donde los niños lo invitaron borde del mar a tirar piedras, comimos (no agregaré adjetivos) en las pescaderías del puerto de San Antonio, donde com-

pró unos mariscos de forma, nombre y color sospechoso, y *bey blades* a los niños, que resultaron ser “truchos”. Lo que quedará, sin embargo, fue el encuentro casual con un contrabandista que en plena calle nos ofrecía un notebook *Toshiba* “livianito, en su caja y recién desembarcado”. Alberto, se enfrascó preguntándole por especificaciones técnicas sobre el caché, la velocidad de procesamiento y otras menudencias. Confieso que nosotros estábamos algo nerviosos, pues en cualquier momento podía aparecer una cuchilla o la policía. Alberto parecía estar en una gran tienda de lujo y tranquilamente preguntaba por otros productos... Es así como nuestro *dealer* terminó ofreciéndonos una pantalla plana gigante por 200 dólares. Si la oferta no hubiese ido también acom-



Figura 4: Alberto, Gonzalito y Cecilia

pañada de una caja de *locos*, de esos prohibidos por esa época, sospecho que Alberto estaría aun allí indagando especificaciones... Luego nos comentaba con su habitual entusiasmo que le había parecido muy interesante esta versión callejera de *la biblia junto al calefón*.

Volvimos por Casablanca y pasamos a visitar los viñedos de la zona. Por esas casualidades, caímos en el medio de la fiesta anual de los productores de vino del valle de Casablanca. En medio de los fracs y vestidos largos de seda, de un ambiente irrespirable a latifundistas y mujeres pintarrajeadas, Alberto en bluejeans y zapatillas, con una prestancia interior soberbia, entró como en su casa, y pidió trago para nosotros, bebidas para los niños, y por si fuera poco, pidió que le mostraran las gigantescas bodegas de toneles de acero inoxidable. Volvimos el resto

del camino a Santiago riendo de la aventura. “Deben haber pensado que era uno de esos dueños viñas algo estrafalario” comentaba. Paramos en el camino a tomar café en *Millahue*, donde Alberto pidió lo único que no conocía del menú: unos pasteles de alcayota, que por supuesto encontró fantásticos.

El viaje a la cordillera días después tuvo ribetes similares. “Hola Alberto!” lo saludó Gonzalito cuando lo pasamos a buscar al hotel. “Hola, cómo han funcionado los *bey blades*?” les replicó Alberto como si tuviese también cuatro años, y conversaron largamente sobre las propiedades del trompito ése. La cordillera y el clima estaban espectaculares. Fuimos directamente a Valle Nevado, donde almorzamos, luego a la Parva, y luego bajamos en Farellones a estirar las piernas y tirar piedras a la quebrada. La competencia la ganó Alberto, ante las justificadas protes-



Figura 5: Alberto con un *quillay* de fondo en una foto que pidió especialmente

tas de trampa de los niños...

Ese fue el último paseo con Alberto en Chile. El viernes 30 de Abril se despidió de todos nosotros: “¿Podríamos salir a cenar este viernes con vuestras respectivas cónyuges?” invitó, con un tono cómplice en aquello de cónyuges. Fue una comida muy agradable. Llegamos Ricardo y yo con nuestras respectivas “cónyuges”, y hablamos monotématicamente de los problemas de nuestros hijos pequeños. Y como para incorporar a Alberto, que disfrutaba la comida peruana escuchando respetuosamente la aburrida monserga, alguien le dijo: “Por suerte tu tienes hijos mayores ya.”

–Hijos pequeños, problemas pequeños. Hijos grandes, problemas grandes, sentenció con sabiduría, y pedagógicamente ilustró su afirmación:

–Para algún viaje, recuerdo, le pedí expresamente a mis hijos que no llevaran a mi pieza a sus novias. Un día me entero que mi cama era un desastre. Sin embargo, nadie había violado la regla: uno de ellos había hecho un festín en mi cama con una dama que no era su novia...

Nosotros con Carlos continuamos comunicándonos con Alberto por correo electrónico. Nos codificó como C&C o c+c. Estructuramos algunos papers sobre nuestras discusiones en Chile. Para siempre nos

## Foundations of Semantic Web Databases

Claudio Gutierrez<sup>1</sup> Carlos Hurtado<sup>2</sup> Alberto O. Mendelzon<sup>1\*</sup>  
<sup>1</sup>Department of Computer Science, Universidad de Chile  
{cgutierr, churtado}@dcc.uchile.cl  
<sup>2</sup>Department of Computer Science, University of Toronto  
amend@cs.toronto.edu

### ABSTRACT

The Semantic Web is based on the idea of adding more machine-readable semantics to web information via annotations written in a language called the Resource Description Framework (RDF). RDF resembles a subset of binary first-order logic including the ability to refer to anonymous objects. Its extended version, RDFS, supports reification, typing and inheritance. These features introduce new challenges into the formal study of sets of RDF/RDFS statements and languages for querying them. Although several such query languages have been proposed, there has been little work on foundational aspects. We investigate these, including computational aspects of testing entailment and redundancy. We propose a query language with well-defined semantics and study the complexity of query processing, query containment, and simplification of answers.

### 1. INTRODUCTION

The Web is a huge collection of interconnected data. Managing and processing such information is difficult due to the fact that the Web lacks semantic information. The Semantic Web is a proposal to build an infrastructure of machine-readable semantics for the data on the Web. In 1998 the W3C issued a recommendation of a metadata model and language to serve as the basis for such infrastructure, the *Resource Description Framework (RDF)* [22]. As RDF evolves, it is increasingly gaining attraction from both researchers and practitioners, and is being implemented in world-wide initiatives such as the Open Directory Project, Dublin Core, FOAF, and RSS.

RDF follows the W3C design principles of interoperability, extensibility, evolution and decentralization. Particularly, the RDF model was designed with the following goals: simple data model; formal semantics and prov-

able inference; extensible URI-based vocabulary; allowing anyone to make statements about any resource. In the RDF model, the universe to be modeled is a set of *resources*, essentially anything that can have a *universal resource identifier*, URI. The language to describe them is a set of *properties*, technically binary predicates. Descriptions are *statements* very much in the subject-predicate-object structure, where predicate and object are resources or strings. Both subject and object can be anonymous objects, known as *blank nodes*. The subject or object of an RDF statement can be another statement, a feature known as *reification*. In addition, the RDF specification includes a built-in vocabulary with a normative semantics (RDFS). This vocabulary deals with inheritances of classes and properties, as well as typing, among other features [24]. Good introductory references for the RDF model are [25] and [26]. Figure 1 shows an example of RDF data.

### 1.1 Problem Statement

Languages for querying RDF have been developed in parallel with RDF itself, e.g. rdqDB [10], SquishQL [15], RQL [13], Triple [19] qREL [16], DQL [21], SoRQL [4]. However, there is very little research so far on foundational aspects of the RDF data model and RDF query languages. Research on formal aspects of RDF data and query languages is made necessary by the new features that arise in querying RDF graphs as opposed to standard databases: the presence of blank nodes, reification, premises in queries, and the RDFS vocabulary with pre-defined semantics. Formal aspects that need study are normal forms and redundancy elimination, entailment, semantics of query languages, query containment and complexity of query processing.

The RDF data model allows several representations for the same information. This raises the question about the existence of normal forms and testing of equivalence among them. Currently different RDF query languages implement different querying mechanisms and functionalities that have not been the subject of a systematic and integrated study. Traditional database notions of query containment do not translate directly to the RDF setting. They need to be reformulated to take into account the fact that RDF queries process logical specifications rather than plain data. Additionally, premises and constraints on queries add further com-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.  
PODS 2004 June 14-16, 2004, Paris, France.  
Copyright 2004 ACM 1-58113-858-X/04/06...\$5.00.

Figura 6: Ultimo paper de Alberto en PODS, iniciado durante su estancia en Chile

quedará la imagen de esos *streams* de mails que de repente comenzaban a llegar cuando le tocaba a Alberto revisar una versión. Terrible, temible, imparabile: línea por línea, demostraciones, cotas de complejidad, gramática, estilo, notaciones, todo era inspeccionado con un rigor excesivo para los usos de nuestra latitud. Generalmente partía inocentemente: **–Un par de preguntas ingenuas para facilitarme la lectura... y luego venía el vendaval.**

En esta comunicación por *email*, conocí diversas facetas académicas de Alberto. El rigor, la capacidad de trabajo, la concentración en un objetivo. Algunas anécdotas que recojo de los *logs* de mi correo electrónico.

En algún momento, ironizando sobre una idea en paralela que le enviábamos y pedíamos que revise: **–OK, no produzcan tantos papers juntos, que no me da el tiempo para leerlos! Le voy a**

**dar prioridad, obviamente, al de RDF.**

Con Alberto sentíamos que teníamos un protector natural en el a veces hostil ambiente científico. En una ocasión, enviamos desde Chile la versión final de un paper del cual no llegó nunca un *acknowledgement*. Peor, el editor reclamó por no haberla enviado. Acudimos a Alberto, que con una breve y sutil mezcla de autoridad y buenas maneras intercedió inmediatamente:

**–As far as I know, the abstract and paper were already sent.**

**–Alberto, thanks so much for acting so quickly!**, fue la respuesta casi automática, y milagrosamente apareció el documento perdido...

Su capacidad de trabajo y sentido de responsabilidad era asombroso. Con un deadline encima, después de intercambiar decenas de mails toda la tarde, y dejar inconclusa la discusión sobre *premisas*

en lenguajes de consultas, nos avisa:

–Voy a estar desconectado hasta última hora, espero poder responder a mails por la noche.

Pero no respondió por la noche. A la mañana siguiente, con encabezado Fri, 28 Nov 2003 06:18:49, nos encontramos con el siguiente mensaje:

–Me acordé de una referencia sobre premises: Hypothetical Datalog, Bonner, PODS 80 y algo.

Para el último paper, tuvimos una memorable discusión sobre la complejidad de determinada operación de reducción en cierto tipo de grafos. Carlos estaba en Grecia, así que me tocó sufrir solo el *stream*. Le envié un ejemplo genérico para sugerir que podíamos encontrar un algoritmo polinomial. No se la tragó. Por el contrario, me convenció que era *NP-hard*, refutándome el ejemplo. “De acuerdo, parece que tu intuición no falla” le escribí. Minutos después recibí otro mail, y otro ejemplo:

–Sin embargo ahora me inclino a creer que es polinómico.

Los papeles se invertían. Me vino a la mente el ajedrez, Alekhine y aquella famosa partida, donde su adversario, jugando con negras, se siente derrotado y se rinde; Alekhine le ofrece continuar la misma partida invirtiendo los roles, le muestra que había una salida, y cuando el adversario se siente derrotado nuevamente, vuelve a hacer lo mismo. Y así un par de veces. Refuté el ejemplo, esta vez ya temeroso de lo que el maestro se traía:

–De acuerdo. Olvidemos ese ejemplo. ¿Tienes una demostración de que la construcción funciona?

El asunto se ponía serio, y ya no bastaban los ejemplos. Envié mi demostración. Refutada. Y él a su vez delineó otra demostración. Esta vez se la refuté yo. Entremedio su salud comenzaba a ser un tema. Ya sabíamos que no era un problema menor. Me atreví tímidamente a preguntarle por su tratamiento. Como siempre, muy reservado, siguió la discusión técnica como si no hubiese preguntado nada. Dos horas y varias decenas de mails más tarde, logré mostrarle por enésima vez el error en la demostración que proponía y sucesivamente parchaba, e hidalgamente escribe bromeando:

–Correcto. ¿Por qué mejor no hablamos de mi salud?

Era el 12 marzo de 2004. Antes, le había contado que Cecilia me había reprochado por ser tan “cuadrado” y de haber intercambiado decenas de mails técnicos sin haberle preguntado por la salud. Alberto respondió:

–No te preocupes, no hay mejor técnica para distraerme de mi salud y mi tratamiento que conversar de teoremas...

Y entonces seguimos indefinidamente conversando de teoremas. Siempre estaba al otro lado del correo electrónico para responder cada pregunta, por absurda que fuera. Para aconsejarnos, para orientarnos. Días antes de agravarse, le enviamos un esquema de *survey* sobre bases de datos de grafos para que nos dé orientaciones estratégicas. Aun en esos difíciles días su respuesta no demoró:

–Cuando tenga tiempo de mirar el survey les contesto. Efectivamente es un tema pasado por alto ultimamente.

Saludos

–a.

Fueron para nosotros los últimos bits del *stream*.

Lo echamos de menos, todos. En el DCC perdimos al gran amigo, quien siempre mostró su inmediata disponibilidad a apoyar a la comunidad en Chile. Cecilia, al caballero, sencillo y de buen carácter. Carlos y nuestro grupo, al maestro y al mentor. Nuestro amigo sopaipillero, a su inusual cliente.

Papá, ¿qué escribes? me preguntó Gonzalito cuando terminaba estas líneas. Sobre Alberto, le respondí. La respuesta fue automática:

–¿Y cuando viene de nuevo?