



Estudiantes DCC



En esta sección de la Revista estudiantes recientemente graduados del Departamento de Ciencias de la Computación (Universidad de Chile) nos cuentan, junto a sus profesores guías, sobre sus trabajos de memoria y/o tesis.



Tesis de doctorado

A Structural Analysis of Dictionaries as Semantic Networks

Estudiante: Camilo Garrido

Profesor guía: Claudio Gutiérrez



Después de titularme de ingeniero civil en computación en la Universidad de Chile y trabajar dos años en la industria, principalmente como ingeniero de software, no lograba sentirme cómodo siguiendo ese camino en el futuro. Buscando alternativas fui a conversar con Claudio Gutiérrez, quién fue mi profesor guía en el pregrado. Le mencioné mi intención de hacer un postgrado y me contó las diferencias entre un magíster y un doctorado, entre estudiar en Chile o el extranjero, las proyecciones y oportunidades después de terminar, entre otros datos que podrían ayudarme en mi decisión. Finalmente, sin estar del todo seguro, decidí aventurarme en hacer un doctorado.

Mi investigación se enfocó en el estudio de las redes de diccionarios lingüísticos. En los diccionarios, por ejemplo el *Diccionario de la Lengua Española* de la RAE, la definición de una palabra utiliza otras palabras y, recursivamente, estas otras palabras utilizan otras más para definirse. De esta manera, un diccionario puede ser visto naturalmente como una red. El objetivo principal fue analizar esta estructura de red para respaldar la hipótesis que las redes léxicas subyacentes son buenas fuentes de material para estudiar lenguaje natural y obtener valiosos resultados. Este trabajo apunta a formar parte de la base de cualquier teoría estructural de diccionarios y esperamos que este tipo de enfoque, como se sugiere en algunos ejemplos, ayude a los lingüistas a estudiar la relación semántica de las palabras. Es una guía de cómo abordar este tipo de redes: describe algunas de sus propiedades comunes, cómo se miden y por qué se cree que son importantes. Entre las características más relevantes de los diccionarios está que comparten una estructura de red común y que es distinta a otros tipos de redes como redes sociales, biológicas u otras. Además, analizando la evolución del dic-

cionario español a través de los años, desde 1925 al 2014, observamos que la incorporación y eliminación de palabras no afectan la estructura de la red. Los cambios son detectables en pequeñas vecindades, dándole a cada palabra una red de significados de sutil evolución. Estos dos puntos muestran que si en un diccionario reemplazamos las palabras por otros símbolos, podríamos darnos cuenta de que lo que estamos viendo es un diccionario pero no podríamos determinar cuál es el idioma original y menos el año de publicación.

Hacer un doctorado fue una gran experiencia y un vaivén de emociones que me permitió aprender mucho. Obviamente uno aprende mucho del área de investigación, pero lo que más rescató de la experiencia fueron las habilidades transversales que fui desarrollando en el camino. Por ejemplo, manejar mis tiempos y aprender a focalizar esfuerzos. Siempre van saliendo nuevas ideas que muchas veces hay que dejar de lado para poder avanzar. Otro ejemplo es lidiar con el rechazo y mantener la motivación. Seguir trabajando en un paper que te rechazaron después de estar meses preparándolo es duro. Como han dicho otros compañeros en esta revista, transmitir tus ideas y presentarlas de forma concisa y clara no es fácil. Es algo que se va ejercitando durante todo el doctorado, ya sea de forma escrita en un paper o de forma oral en una presentación.

Luego de haber terminado el doctorado, decidí no seguir un camino académico, pero no por ello alejado de la investigación. Ingresé a trabajar en la Dirección de Innovación y Transferencia del Instituto Milenio Fundamentos de los Datos, cuyo objetivo es generar y fortalecer lazos con organizaciones públicas y privadas que tengan requerimientos urgentes en materias relacionadas con datos.



Tesis de magister



Detección temprana de anomalías de un proveedor de plataformas de telecomunicaciones

Estudiante: Juan Andrés Calvo Rodríguez

Profesor guía: Nelson Baloian

Las herramientas de monitoreo son un gran apoyo a los equipos de Operaciones, ya que permiten vigilar la salud del sistema, detectando anomalías y realizando acciones que sin ellos solo dependerían del reclamo de algún cliente final. Cuando se producen reclamos por parte de clientes finales por algún producto o servicio, generalmente el problema que los provoca viene afectando la operación comercial por un tiempo considerable, lo que se traduce en una mala evaluación de la compañía de telecomunicaciones. Esto deriva en pérdida de tiempo y dinero para corregir la anomalía, arreglar información distorsionada por el problema, y por último también realizar acciones para recuperar la buena percepción de la compañía. Para evitar esto, en una compañía que brinda servicios de telecomunicaciones se diseñó una herramienta de monitoreo, de manera de alertar a un especialista para que revise la anomalía y realice las acciones correctivas o paliativas con anticipación.

Primero se revisaron implementaciones de monitoreo en el mercado, se realizaron entrevistas con especialistas y usuarios. El diseño incorporó indicadores necesarios para tener una idea de la salud general del sistema, muchos de estos indicado-

res sugeridos por los especialistas, que calculan manualmente cuando se reporta una incidencia y otros fueron considerados a partir de otras soluciones revisadas, como por ejemplo los indicadores de los servidores (CPU, ocupación de memoria, etc.). En el análisis del sistema se detectó concentración de datos en la dispersión de duración por cantidad de transacciones, lo que permitió verificar que es posible utilizar la distribución normal. Modelando los datos se observó concentración de valores no anómalos y también se advirtió que los datos anómalos se sitúan en el extremo de la gráfica. Se recolectaron datos y casos de incidencias (anomalías), algunas reportadas por los usuarios y otras que fueron encontradas por el Equipo de Soporte durante la preparación de informes cuando se buscó la causa raíz de un problema.

Se realizó análisis de estos datos calculando los indicadores de manera manual para validar que es posible encontrar anomalías a partir de dichos indicadores, se seleccionaron casos y luego, se implementó el algoritmo de detección temprana basado en distribución normal utilizando dos variables: cantidad de transacciones y duración promedio de dichas transacciones, y se simuló el comportamiento del sistema utilizando los respectivos grupos de datos, comparando resultados hallados manualmente y resultados generados por el algoritmo. Se observó que la implementación del algoritmo es capaz de detectar la primera anomalía del día, es decir alertar la anomalía lo antes posible (al menos doce horas antes que lo reporte una persona), de manera que el especialista cuente con más tiempo para analizar y realizar acciones de mitigación. La solución de monitoreo también contempló el diseño de una aplicación web, para mostrar resultados del monitoreo y acelerar el análisis; esta levantará alarmas cuando alguno de los indicadores pase el umbral de valores válidos, que estarán configurados, y también cuando lo informe el algoritmo de detección temprana.



Tesis de magister



Reconocimiento de entidades anidadas en diagnósticos de la lista de espera en hospitales públicos

Estudiante: Matías Rojas
Profesores guías: Felipe Bravo Márquez y Jocelyn Dunstan

El sistema de salud en Chile está dividido en dos grupos principales, el sistema público (FONASA) y el sistema privado (ISAPRES). En el sistema público, cuando un paciente quiere visitar a un médico especialista, primero debe acudir a una interconsulta con un médico de atención primaria. En esa interconsulta, se define si ese paciente debe o no ser derivado al especialista, lo que se maneja a través de la llamada lista de espera. Dada la alta demanda por especialistas, este sistema ha alcanzado un tiempo de espera promedio de más de un año, convirtiéndose en un problema importante a nivel nacional. En este contexto, la investigación realizada por Matías en conjun-

to con los profesores Felipe Bravo Márquez y Jocelyn Dunstan, propone lo siguiente:

Cuando el médico de atención primaria atiende a un paciente, se genera un texto no estructurado que contiene la sospecha diagnóstica, y que gracias a la ley de transparencia, se puede acceder públicamente. En estos textos se pueden encontrar palabras claves asociadas a enfermedades, medicamentos, hallazgos clínicos u otra información con relevancia clínica, y que debido a su naturaleza no estructurada, han sido poco estudiados. Dado esto, se desarrolló un modelo basado en inteligencia artificial que es capaz de reconocer automáticamente esta información, labor que no es sencilla incluso para los médicos especialistas. La creación de estos modelos permite, entre otras cosas, generar estudios epidemiológicos, determinar si alguna enfermedad puede ser tratada con telemedicina, obtener estadísticas sobre la lista de espera, o simplemente acelerar el proceso mediante el cual se analizan estos diagnósticos, considerando la alta demanda.

Este trabajo y sus aplicaciones han contribuido a la publicación de artículos científicos, el desarrollo de una plataforma para probar el modelo y la obtención del premio a los 100 Jóvenes Líderes por parte de la revista Sábado, destacando la importancia de esta investigación en el contexto nacional.

Memoria de pregrado



Gesture Coding: Diseño de un entorno de programación interactiva controlado por Joy-Cons

Estudiante: Emilio Lizama Castro
Profesor guía: Francisco J. Gutiérrez

Hoy en día, somos testigos de un auge en la exploración de nuevas formas para desarrollar el pensamiento computacional en niñas y niños. Tradicionalmente, el mecanismo preferido para acercar a los jóvenes a la programación, sobre todo a aquellos/as en edades tempranas (10-12 años), es a través de lenguajes de programación basados en bloques, como Scratch, Blockly, Snap! o App Inventor. La mayoría de estos lenguajes, si es que no todos, se basan en un paradigma de interacción tradicional de tipo *point-and-click* y escribiendo comandos e instrucciones con un teclado.

Aun cuando este tipo de mecanismos de entrada no representan problemas significativos de usabilidad en su adopción, reconociendo la diversidad en los modos de interacción contemporáneos vemos en este trabajo una oportunidad en explorar mecanismos novedosos para despertar el interés en desarrollar el pensamiento computacional y la programación en escolares. Para esto, se diseñó y desarrolló "Gesture Coding", una biblioteca de metáforas de interacción natural (es decir, gestos) junto a un entorno de desarrollo integrado (IDE). Estos dos utilizan como dispositivo de entrada Joy-Cons, controles de movimiento de la consola Nintendo Switch. En particular, con esta herramienta esperamos que controles de entrada interactivos, lúdicos y no tradicionales, reduzcan la barrera de entrada para un uso continuo, facilitando así la aceptación y apropiación del entorno de programación.

Para validar la usabilidad y funcionalidad del IDE desarrollado, se ejecutó una prueba de concepto. Los resultados mostraron que Gesture Coding es una alternativa novedosa, funcional, usable y considerada como divertida por parte de niñas y niños. Este trabajo presenta un primer acercamiento a diseñar un mecanismo que utilice metáforas de interacción natural para la enseñanza del pensamiento computacional en escolares.



Memoria de pregrado



Herramienta para análisis de programas desarrollados en Scratch

Estudiante: María José Berger
Profesora guía: Jocelyn Simmonds

El fomento del pensamiento computacional en niñas y niños se ha convertido en un tópico de interés mundial, dado los procesos cognitivos que promueve en las y los estudiantes. Las habilidades que desarrolla el pensamiento computacional refuerzan habilidades orientadas a la resolución de problemas, como el pensamiento crítico, la creatividad y la colaboración. En esta línea, desde el 2012, el grupo REACT (Rethinking Education by Advancing Computational Thinking) del DCC ha estado explorando estrategias y metodologías para introducir el pensamiento computacional en el aula.

Por ejemplo, REACT realiza talleres de desarrollo del pensamiento computacional para niños, niñas, adolescentes, profe-

sores y profesoras. En estos talleres, los participantes siguen actividades prácticas en las que adquieren los conceptos básicos de programación y desarrollan proyectos de software a pequeña escala utilizando Scratch. Si bien existen herramientas como Dr. Scratch, que permite realizar análisis de frecuencia del uso de instrucciones como una forma de evaluar las competencias logradas por el/la autor/a de un programa creado en Scratch, estas herramientas no incorporan la estructura de los programas en su análisis.

En este trabajo de título, se creó una plataforma para el análisis de programas en Scratch basado en su estructura. Dado que Scratch no posee especificación formal, esta plataforma también implementa un *parser*, lo que permite trabajar directamente con el Diagrama de Control de Flujo de un programa, en vez de su especificación en Scratch. Este diagrama representa todas las posibles vías de computación en un programa, y permite usar prácticas tradicionales de navegación de grafos y consultas de bases de datos para ejecutar análisis de programas en base a su estructura. La plataforma se implementó como una aplicación Web, que expone una API y una interfaz donde se cargan programas individual o masivamente para su análisis a enseñanza del pensamiento computacional en escolares.