

Primera Jornada de Doctorandos de Computación FaMAF, Universidad Nacional de Córdoba

6 de Diciembre, 2010

Programa del Evento

	09.00 - 09.10	Apertura
	09.10 - 09.30	<i>Raul Fervari</i> Lógicas Modales: Clásicas y Dinámicas
	09.30 - 09.50	<i>Romina Altamirano</i> Mejorando la traducción automática a través de la incorporación de conocimiento lingüístico, simplificación textual, y combinación de motores de traducción
	09.50 - 10.10	<i>Araceli Acosta</i> Herramientas y mecanismos formales para el tratamiento de la tolerancia a fallas: Síntesis de controladores
	10.10 - 10.30	<i>Miguel Pagano</i> Normalización por evaluación
	10.30 - 11.00	Coffee Break
	11.00 - 11.20	<i>Martin Dominguez</i> Análisis no supervisado de dependencias basado en Gramáticas Biléxicas
	11.20 - 11.40	<i>Hernan Casalánguida</i> Desarrollo de Familias de Aplicaciones Web con Transformación de Modelos
	11.40 - 12.00	<i>Julio Castillo</i> Reconocimiento de Implicación Textual empleando Aprendizaje Automático
	12.00 - 12.20	<i>Eduardo Romero</i> AI Planning and Scheduling: enfoques, formalismos y aplicaciones
	12.20 - 14.00	Almuerzo

	14.00 - 14.20	<i>Franco Luque</i> Inducción No Supervisada de Gramáticas de Lenguaje Natural
	14.20 - 14.40	<i>Cecilia Kilmurray</i> Utilizando Lógicas Deónticas y Temporales para la Especificación de Sistemas Tolerantes a Fallas
	14.40 - 15.00	Charla 11: <i>Renzo Degiovanni</i> Análisis Automático basado en SAT aplicado a la Validación de Requisitos de Software
	15.00 - 15.20	<i>Valeria Bengolea</i> Mejoras en la Generación de casos de Test en Korat usando Criterios de Test
	15.20 - 16.30	Merienda
	16.30 - 16.50	<i>Nicolás Wolovick</i> Measurability for Safety Verification of Stochastic Hybrid Systems
	16.50 - 17.10	<i>Ezequiel Orbe</i> Machine-based Heuristics for propositional SAT solving
	17.10 - 17.30	<i>Matias Lee</i> A Refinement Based Notion of Non-Interference for Interface Automata: Compositionality, Decidability and Synthesis
	17.30 - 17.50	<i>Silvia Pelozo</i> Razonando sobre seguridad en sistemas distribuidos probabilistas por medio de alcanzabilidad acotada
	17.50 - 18.10	<i>Renato Cherini</i> Verificación de sistemas de código que utilizan memoria dinámica
	18.10 - 18.30	Cierre

Resumen de las Charlas

- Araceli Acosta [*acostaaraceli AT gmail.com*]

Herramientas y mecanismos formales para el tratamiento de la tolerancia a fallas: Síntesis de controladores

Contaré un poco de qué se trata mi tema de doctorado, y una pequeña introducción a la técnica formal con la que estoy trabajando en este momento, "Síntesis de controladores", y su aplicación a la tolerancia a fallas.

- Romina Altamirano [*romina.altamirano AT gmail.com*]

Mejorando la traducción automática a través de la incorporación de conocimiento lingüístico, simplificación textual, y combinación de motores de traducción

La traducción automática es una tarea compleja porque requiere análisis de texto de origen, resolver ambigüedades de diversos tipos y generación de texto en el idioma destino. Actualmente no existe un traductor de propósito general que exitosamente haga su tarea en todos los idiomas y dominios, Mejorar algún aspecto de la traducción automática reduciría notablemente el trabajo de post-edición, entre esas mejoras que se pueden realizar están la de agregar conocimiento lingüístico, simplificar el texto en una etapa de pre-edición, y combinar motores de traducción para explotar mejor las ventajas de cada uno de ellos.

- Valeria Bengolea [*vbengolea AT dc.exa.unrc.edu.ar*]

Mejoras en la Generación de casos de Test en Korat usando Criterios de Test

La generación automática de casos de test es una área de investigación muy activa en la actualidad, dentro de la cual la generación de casos de test para estructuras complejas alojadas en memoria dinámica, tales como árboles balanceados, grafos, etc., es particularmente difícil. Korat es una herramienta/algorithm que resuelve este problema de manera bastante efectiva, mediante la exploración exhaustiva de potenciales casos de test. El mecanismo de constraint solving de Korat está basado en búsqueda (backtracking), y depende fuertemente de una poderosa técnica de poda. En esta charla presentaré una técnica para mejorar la generación de casos de test de la manera realizada por Korat, guiando la búsqueda a través de la consideración de un criterio de test. Tal criterio se emplea para poder podar la búsqueda aún más, pero de una manera similar a la utilizada por Korat. Presentaré la técnica junto con algunos resultados experimentales que muestran la performance de la técnica desarrollada para algunos criterios de test, y una comparación con Korat estándar.

- Hernán Casalánguida [*9dearea AT gmail.com*]

Desarrollo de Familias de Aplicaciones Web con Transformación de Modelos

El problema general a estudiar es el de definir una metodología apropiada para el desarrollo de familias de aplicaciones web (FAW) intensivas en datos e incluyendo aplicaciones de internet ricas. Actualmente existe una presión muy grande para disminuir los costos y los tiempos de desarrollo de aplicaciones web; al mismo tiempo los diversos atributos de calidad de este tipo de software se ven afectados por el estilo de desarrollo desde cero. La disponibilidad de métodos, conceptos de modelado, abstracciones básicas y por supuesto frameworks de desarrollo permitiría contribuir de manera muy importante a esta problemática. La ingeniería de FAW permitirá a las empresas mayores utilidades, mayor calidad en los productos y tiempos más cortos de desarrollo al poder construir las aplicaciones web de una FAW apoyándose en una infraestructura reutilizable. A pesar que la ingeniería de familias de aplicaciones (IFA) es un área que empezó al comienzo de los 80, casi no se ha estudiado la ingeniería de FAW y la investigación permitirá mejorar y comprender mejor esta nueva rama de la IFA.

- Julio Castillo [*jotacastillo AT gmail.com*]

Reconocimiento de Implicación Textual empleando Aprendizaje Automático

Dados dos fragmentos de textos llamados T(Texto) e H(Hipótesis), la tarea de reconocimiento de implicación textual consiste en determinar si el significado de H puede ser inferido tomando T como válido o en el contexto de T. De esta manera, la noción de "implicación" se define como una relación unidireccional que se sostiene entre dos expresiones del lenguaje, mientras que la relación opuesta podría no mantenerse. Como resultado podríamos tener que esta relación se sostiene, que la hipótesis H contradice el texto T, o bien que no hay suficiente información o se desconoce tal relación. Así, en esta charla se dará una introducción al problema del Reconocimiento de Implicación Textual, y a una aproximación utilizando aprendizaje automático para su resolución. Se hablará sobre la evolución de dicha tarea desde su definición inicial y se comentarán algunos resultados obtenidos hasta el momento.

- Renato Cherini [*cherini AT gmail.com*]

Verificación de sistemas de código que utilizan memoria dinámica

La Separation Logic es el estándar de facto en la verificación formal de programas que utilizan memoria dinámica para implementar estructuras de datos mutables. Se introducirá brevemente esta lógica y se discutirá sobre dos limitaciones que se encuentran al utilizarla en la verificación de sistemas de código: por un lado, en cuanto a la expresividad, la imposibilidad de especificar estructuras de datos que comparten la memoria para su representación de forma compatible con el principio de "information hiding" característico de los actuales paradigmas de programación imperativa; por otro lado, en un aspecto mas práctico, la complejidad de las demostraciones necesarias y la dificultad de llevarlas adelante de forma manual. Se presentarán dos trabajos que apuntan a resolver estas dificultades.

- Renzo Degiovanni [*rdegiovanni8 AT gmail.com*]

Análisis Automático basado en SAT aplicado a la Validación de Requisitos de Software

Existe una amplia variedad de notaciones para la descripción de requisitos de software, pero aquellas con una semántica formal son particularmente apropiadas para el análisis. En particular, la notación tabular de Parnas y el método Software Cost Reduction (SCR) resultan convenientes para la descripción formal de requisitos, pues basan sus notaciones en formas adecuadas de organizar expresiones relacionales, de manera tal que su comprensión y manipulación sea más efectiva. En esta charla discutiremos algunas ideas de cómo atacar el problema de utilizar análisis basado en SAT para tareas ligadas a la validación de requisitos de software especificados en las notaciones tabulares mencionadas, y aprovechando características propias de estas notaciones. Entre los objetivos de análisis se encuentran el conseguir documentos de requisitos completos y consistentes. Veremos que el análisis basado en SAT requiere, al igual que otras técnicas de análisis automático, de mecanismos que contribuyan a su escalabilidad. Discutiremos uno particular, basado en abstracción, que resulta útil en el contexto de especificaciones tabulares.

- Martin Dominguez [*mardom75 AT gmail.com*]

Análisis de Dependencias no supervisado de dependencias basado en Gramáticas Biléxicas

El análisis de dependencias no supervisado, cuyo objetivo es adquirir conocimiento de un lenguaje a partir de un corpus sin ninguna anotación, ha tomado gran importancia en el área del procesamiento natural en los últimos años. Es utilizado como parte de sistemas de traducción automática, sistemas de diálogos, buscadores y respuesta a preguntas (question answering), entre otros. El análisis de dependencias tiene como foco construir un modelo que permita inferir el árbol de dependencias para una oración dada. En esta charla introducimos un analizador de dependencias no supervisado basado en gramáticas biléxicas. El modelo está basado en la inferencia de una gramática bilexica, la cual se construye a partir de un conjunto de autómatas que modelan el lenguaje de dependencias para cada categoría sintáctica (verbos, sustantivos, etc). Los resultados obtenidos por nuestro analizador son comparables con el estado del arte para el idioma inglés en oraciones de hasta diez palabras. La arquitectura definida, al estar basada en autómatas, permite mucha flexibilidad; permitiendo definir distintas estructuras, dependiendo de los idiomas y de las categorías sintácticas. Debido a esto, esperamos en trabajos futuros obtener buenos resultados tanto para oraciones de mayor longitud,

como para otros idiomas como el alemán, el turco y el español.

- Raúl Fervari [*rfervari AT gmail.com*]

Lógicas Modales: Clásicas y Dinámicas

En la charla se presentará una introducción a algunos lenguajes lógicos simples para mostrar la importancia de este tipo de lenguajes en nuestro trabajo (aplicaciones, preguntas a responder, resultados generales, etc.), y sus limitaciones (expresividad, complejidad, etc.). Luego se procederá a introducir de una manera más formal el lenguaje modal básico, qué nos permite decir acerca de estructuras relacionales, y se mostrará por qué no resulta suficiente para explorar todas las características de dichas estructuras, haciendo referencia a propiedades dinámicas de los modelos. Para estudiar estas ideas, se presentarán dos lenguajes creados con este fin: *memory logic* y *swap logic*. El primer lenguaje se introducirá formal e informalmente a modo de ejemplo, mientras que el segundo, por ser específicamente el tema en el que estamos trabajando en el momento, se presentará mediante algunos resultados preliminares y discutiendo problemas aún sin resolver.

- Cecilia Kilmurray [*ckilmurray AT dc.exa.unrc.edu.ar*]

Utilizando Lógicas Deónticas y Temporales para la Especificación de Sistemas Tolerantes a Fallas

En la actualidad la tolerancia a fallas cada vez adquiere más importancia, esto se debe a que cada día hay más sistemas críticos en donde es necesario garantizar cierto comportamiento deseado aún ante la ocurrencia ocasional de fallas. Ejemplos de estos sistemas son: centrales nucleares, aviones, dispositivos médicos y cualquier otro sistema que se caracterice por un funcionamiento continuo, durante largos plazos de tiempo, y sin posibilidad de intervención. El interés en este tipo de sistemas, con altos niveles de disponibilidad y confiabilidad, ha llevado a un gran desarrollo en esta área de la computación. Sin embargo una característica sorprendente de la tolerancia a fallas es la falta de soluciones de propósito general, o ambientes para el análisis de sistemas tolerantes a fallas. La amplia mayoría de las soluciones a problemas de tolerancia a fallas son *ad hoc*, y atacan problemas particulares en contextos específicos. En el caso de técnicas para tolerancia a fallas en software, las soluciones más importantes tienen que ver con redundancia o diversidad. En esta charla presentaré brevemente las características básicas de algunos frameworks lógicos que pensamos utilizar para razonar sobre tolerancia a fallas. En particular mostraré algunos ejemplos para ilustrar la capacidad de ciertas lógicas deónticas para capturar nociones relacionadas con tolerancia a fallas. Culminaré presentando algunas de las ideas de lo que respecta a análisis probabilístico de propiedades de sistemas, como esto se relaciona con lo antes explicado y el plan de trabajo a futuro.

- Matias Lee [*lee.matias AT gmail.com*]

A Refinement Based Notion of Non-Interference for Interface Automata: Compositionality, Decidability and Synthesis

Interface automata (IA) introduce a framework to model stateful interfaces. Interface structures for security (ISS) extend IA to cope with security properties. In this work, we argue that bisimulation-based non interference is not quite appropriate to characterize security on ISSs. We instead introduce a refinement-based variant of non-interference that fits better in this context. Moreover, we show that these new properties are not preserved by composition, but give sufficient conditions to ensure compositionality. We give two algorithms. The first one determines if an ISS satisfies the refinement-based non-interference property. The second one, determines if an ISS can be made secure by controlling some input actions, and if so, synthesizes the secure ISS.

- Franco Luque [*francofq AT gmail.com*]

Inducción No Supervisada de Gramáticas de Lenguaje Natural

Uno de los problemas centrales del área de Procesamiento de Lenguaje Natural es el desarrollo de analizadores sintácticos (parsers) de oraciones. En el caso supervisado, los parsers "aprenden" la gramática de una lengua en una fase previa de entrenamiento, en la que se utiliza un corpus de texto previamente analizado sintácticamente por lingüistas. El caso no supervisado es una versión mucho

más compleja del problema de análisis sintáctico. En este caso, los parsers se entrenan solamente con oraciones sin analizar, es decir, con texto de lenguaje natural en bruto. El parsing no supervisado resulta interesante para su aplicación en idiomas para los que no existen corpus sintácticos, y también para modelar la adquisición del lenguaje por parte de los humanos, un problema central de estudio de las ciencias cognitivas. Un área relacionada al parsing no supervisado de lenguaje natural es el de la Inferencia Gramatical de lenguajes formales. En este ámbito se estudia la posibilidad de aprender un lenguaje formal a partir de una secuencia de ejemplos del lenguaje. Existen definiciones formales para aprendibilidad de diferentes tipos, y resultados de aprendibilidad para diferentes clases de lenguajes dentro de la jerarquía de Chomsky. En esta charla presentaré el problema de parsing no supervisado desde el punto de vista de la Inferencia Gramatical, relacionando el lenguaje natural con los lenguajes formales.

- Ezequiel Orbe [ezequiel.orbe AT gmail.com]

Machine-based Heuristics for propositional SAT solving

This talk, consists of two parts. In the first one, I will present the main goals of our research, which aims at the development of new machine-learning based algorithms and heuristics, and the improvement of the existing ones, in particular, branching heuristics, in the context of complete sat solvers. In the second part, I will present a brief introduction to our current work, namely, exploiting symmetry in propositional sat-solving. The symmetry of a discrete object can be defined informally as a permutation of its components that leaves the object intact. In the context of sat-solving we can defined symmetry as a mapping of a sat problem onto itself that preserves its structure as well as its solutions. Many classes of reasoning problems, in particular, those arising from real world applications, displays a large amount of symmetry. In mathematical and common-sense reasoning, such symmetries are often used to reduce the difficulty of reasoning. The use of symmetries in search problems is conceptually simple. If several points of the search space are related by a symmetry then we never want to visit more than one of them. In order to accomplish this we must solve two problems. First, the symmetries need to be discovered. Second, we need to make use of the symmetries. The talk will present a survey of the theoretical background and techniques that are available to attack this two problems.

- Miguel Pagano [miguel.pagano AT gmail.com]

Normalización por evaluación

Un sistema de tipos tiene además de las reglas de tipado algunas reglas de computación, que determinan una noción de forma normal para los términos. Una manera de obtener esa forma normal es utilizando Normalización por Evaluación; en esta técnica se utiliza un modelo semántico para el cual se puede definir una función de reificación que mapea ciertos elementos del dominio semántico a términos en forma normal. En esta charla presentaremos la técnica para el cálculo lambda simplemente tipado y luego comentaremos sobre trabajos recientes en Normalización por Evaluación para teorías de tipos dependientes.

- Silvia Pelozo [spelozo AT famaf.unc.edu.ar]

Razonando sobre seguridad en sistemas distribuidos probabilistas por medio de alcanzabilidad acotada

La charla presenta avances del trabajo en curso para desarrollar un algoritmo que calcule propiedades cuantitativas de alcanzabilidad en sistemas distribuidos. Consideramos sistemas con comportamiento no determinista y probabilista, cuyos componentes deben basar sus decisiones sólo en información local. Esto es importante por dos razones: primero porque los límites de los valores de probabilidad son más realistas, y segundo porque permite razonar sobre propiedades de seguridad. Las técnicas tradicionales de model-checking probabilista resultan inadecuadas en estos sistemas, porque se basan en todas las posibles resoluciones del no determinismo introducido al intercalar procesos, considerando también opciones poco realistas, como por ejemplo que un proceso elija su siguiente acción de acuerdo a lo observado en otro proceso. Para salvar esta limitación recurrimos a restricciones sobre la resolución del no determinismo, como las de los schedulers distribuidos, los

fuertemente distribuidos, y en particular en esta charla proponemos restricciones adicionales que permiten razonar sobre propiedades de seguridad. Por cuestiones de decidibilidad no es posible calcular propiedades de alcanzabilidad no acotada, pero discutimos un algoritmo para calcular alcanzabilidad acotada, consistente en la interpretación de la resolución del no determinismo como parámetros, con lo que se reduce el problema de model-checking a un problema de optimización no lineal.

- Eduardo Romero [*eromero AT conae.gov.ar*]

AI Planning and Scheduling: enfoques, formalismos y aplicaciones

Los problemas de planificación y scheduling aparecen en contextos de lo más variados, y muchas veces están profundamente relacionados como un único problema. Sin embargo, tanto en la práctica como en la teoría, las técnicas utilizadas para resolverlos divergen profundamente en la forma de representación, en su alcance y en los tipos de problemas que son capaces de abordar. En esta charla se presentará de qué se trata este tipo de problemas y varios enfoques existentes para resolverlos. Pasando a las aplicaciones, voy a comentar cómo aparecen estos problemas de manera natural en las operaciones de misiones espaciales.

- Nicolás Wolovick [*nwolovick AT gmail.com*]

Measurability for Safety Verification of Stochastic Hybrid Systems

Dealing with the interplay of randomness and continuous time is important for the formal verification of many real systems. Considering both facets is especially important for wireless sensor networks, distributed control applications, and many other systems of growing importance. An important traditional design and verification goal for such systems is to ensure that unsafe states can never be reached. In the stochastic setting, this translates to the question whether the probability to reach unsafe states remains tolerable. We consider stochastic hybrid systems where the continuous-time behavior is given by differential equations, as for usual hybrid systems. The discrete jumps are given by either continuous probabilities, or discrete probability distributions with continuous nondeterminism. The latter is exploited in an abstraction method that establishes safe upper bounds on reachability probabilities. To arrive there requires us to express the model in a semantic setting that could handle both continuous probabilism and nondeterminism. We will show that Nondeterministic Labeled Markov Process (NLMP) is an adequate model to capture this semantics. These examples will bring evidence that the hit sigma algebra, one of the novel ingredients of NLMP, is adequate for doing the proofs. We will also explore how to specify probabilistic nondeterminism using the generators $\Delta^P(A)$ of the set of probabilities sigma algebra $\Delta(\Sigma)$.