REACT: Robust theories for Emerging Applications in Concurrency Theory

Grupo de Investigación AVISPA

Pontificia Universidad Javeriana, Cali

URL: http://avispa.puj.edu.co

Email: avispa@googlegroups.com

Noviembre de 2006

Agenda

- Introducción
  - Concurrency Theory / Teoría de la Concurrencia
  - Programación Concurrente por Restricciones (CCP)
  - El lenguaje ntcc
- El proyecto REACT
  - Motivación, Objetivos
  - Grupos e Instituciones Participantes
  - Areas de aplicación
    - Biología Sistémica
    - Interacción Multimedial
    - Seguridad Informática
- Oportunidades

Contexto

Teoría de la Concurrencia (Concurrency Theory)

- Métodos formales (matemáticamente fundamentados) para el desarrollo de software / hardware
- Teorías y métodos para la descripción y análisis de sistemas que exhiben comportamiento concurrente.
- Generalmente, estas técnicas se reunen en forma de *frameworks*
- Una metodología sistemática para la verificación y/o el descubrimiento de propiedades esenciales de sistemas en diversas áreas

Contexto

Un *framework* en concurrencia estudia sistemas y/o problematicas bien definidas. Usualmente incluye:

Componente	Técnica Básica	Capacidad
Lenguajes *abstractos* de especificación	Cálculos de Procesos	DESCRIPCIÓN
Mecanismos que formalizan la *evolución* de los sistemas	Semánticas Operacionales	OBSERVACIÓN
Técnicas de *prueba* para expresar y verificar propiedades	Lógica	VERIFICACIÓN

Contexto

Cálculos de Procesos

- Son lenguajes de programación abstractos con componentes reducidos
  - **Abstraen** aquellos elementos de un sistema concurrente que se consideran fundamentales para el análisis.
  - Son **composicionales:** la especificación de un sistema viene de la especificación de sus sub-sistemas
- Las especificaciones suelen ser sucintas y precisas. Esto es conveniente:
  - En el estudio detallado de la evolución de los sistemas
  - En la definición y verificación de propiedades esenciales
  - En la implementación de herramientas de software como intérpretes y simuladores
- AVISPA tiene experiencia en la definición e implementación de cálculos de procesos

Contexto

Aplicaciones

- Inicialmente, la teoría de la concurrencia se ocupó exclusivamente del estudio de sistemas de cómputo moviles/distribuidos
- Existen similitudes entre las ideas de la teoría de la concurrencia y otras áreas del conocimiento:
  - Biología Sistémica
  - Interacción Multimedial
  - Seguridad Informática (Protocolos)
- Cada área involucra formas de interacción particulares y diferentes tipos de propiedades
- Tendencia: desarrollo de frameworks especificos para cada aplicación (o tipo de aplicación).

Contexto

Programación Concurrente por Restricciones (CCP)

- Un **modelo** de concurrencia basado en información parcial
- Varios cálculos de procesos, *frameworks*, e incluso lenguajes de programación de alto nivel (Mozart-Oz).
- Restricción como elemento básico de *información parcial*:
  - Las variables pueden estar parcialmente instanciadas.

    Ej. X es mayor que 42 pero menor que 60
- Es un modelo declarativo:
  - Lo *qué se debe hacer* en lugar de *cómo se debe hacer*

Contexto

- La memoria acumula información en un **almacén de restricciones**. Operaciones básicas:
  - Tell: agrega una nueva pieza de información
  - Ask: consulta el almacén sobre la presencia de una pieza de información
  - Ask y tells definen la sincronización del sistema
- Un **sistema de restricciones** da coherencia y capacidades de inferencia a la información del almacén
  - Ej. De la restricción X > 8 se deduce que X es mayor que 0.
  - Sobre números enteros, reales, complejos, grafos, ….

Contexto

ntcc es un framework de concurrencia basado en CCP desarrollado por AVISPA

- Restricciones, Tiempo Discreto y No determinismo
- Algunas propiedades expresables en ntcc:
  - “El sistema se ejecutará eventualmente en el futuro”
  - “Si una pieza de información está presente, se garantiza la ejecución infinita del sistema”
  - “Dada una serie de alternativas, el sistema puede ejecutar cualquiera de ellas en el siguiente instante de tiempo discreto”
  - “A menos que se reciba un estímulo determinado, el sistema seguirá una ejecución correcta”
- Ideal para analizar sistemas que interactuan continuamente (incluso infinitamente) con su entorno

Contexto

- ntcc ofrece
  - Un lenguaje de especificación
  - Mecanismos semánticos basados en la observación de los sistemas
  - Sistema de pruebas para verificación de propiedades temporales
- Los sistemas pueden verse, al mismo tiempo, como agentes computacionales y formulas lógicas.
- AVISPA ha implementado intérpretes y máquinas abstractas de procesos ntcc

Contexto

- Hemos utilizado ntcc en diversos tipos de aplicaciones
  - Sistemas en informática musical, en donde hay patrones complejos de temporalidad y sincronización
  - Especificación de robots
  - En el modelamiento y verificación de sistemas biológicos en donde el no determinismo debe combinarse apropiadamente con información parcial
- Hemos encontrado situaciones donde ntcc se queda corto…
  - Patrones temporales irregulares
  - Elementos probabilísticos en la ejecución de sistemas
  - Nombres únicos (esenciales en seguridad)

REACT

El proyecto REACT pretende *fortalecer* los fundamentos teóricos y las herramientas derivadas de ntcc.

Interés específico en tres áreas de aplicación:

- Biología Sistémica
- Seguridad en Protocolos
- Interacción Multimedial

Metodología prevista (por cada área):

- Caracterización de los tipos de interacciones más relevantes y complementación de los formalismos existentes
- Desarrollo de casos de estudio, tomados de sistemas reales
- Diseño e implementación de herramientas de software que permitan la verificación de propiedades sobre los sistemas.

REACT: Participantes

REACT es un esfuerzo internacional entre:

- Grupo AVISPA
- El equipo INRIA Comète, del laboratorio de informática de l'École Polytechnique de París
  - Seguridad, Biología Sistémica
- IRCAM, instituto francés de investigación en acústica y música
  - Interacción Multimedial

REACT apunta a la formación de jóvenes investigadores en las áreas de interés, a nivel de pregrado y postgrado. Cuenta con el respaldo de COLCIENCIAS durante dos años (2007-2008).

Biología Sistémica

De datos aislados a la explicación de funciones biológicas

- Los avances en biología molecular se traducen en una gran cantidad de datos sobre entidades biológicas (genes, proteinas, ADN, ARN)
- Dichos datos deben ser estructurados para explicar el funcionamiento de dichas entidades
- Esta estructuración debe ser a multiples niveles de abstracción
- La idea es “subir” en el nivel de abstracción, refinando progresivamente las hipótesis
- Idea: Modelar entidades biológicas como procesos concurrentes

Biología Sistémica

La información parcial es inherente a la biología sistémica

El descubrimiento de información, desde el punto de vista biologico, es cíclico:

- Se establecen unas hipótesis controvertidas y se introducen en modelos biológicos
- Se simulan dichos modelos utilizando técnicas computacionales (“experimientos secos”)
- Se analizan los resultados de dichos modelos y se reformulan las hipótesis iniciales
- Se realizan experimentos reales, se analizan los resultados, y se reinicia el ciclo

Biología Sistémica

Nuestro "nicho" son los experimentos secos que preceden a los experimentos reales.

Algunas premisas / intuiciones:

- La información parcial puede ser un factor determinante para expresar y aprovechar las hipótesis de trabajo de los biólogos
- La inclusión de información cuantitativa es transparente en CCP
- El comportamiento no determinístico de ntcc, combinado con su manejo temporal, han demostrado ser convenientes en el contexto biológico.
- La estrecha relación entre ntcc y Mozart hace que el desarrollo de herramientas sea inmediato.

Biología Sistémica

- Nuestros intentos iniciales incluyen modelos de:
  - Sistemas de transporte activo en membranas celulares
    - La Bomba de Sodio - Potasio
  - Redes de regulación genética
    - El operón de la lactosa (nivel celular y molecular)
  - Cooperatividad en viruses
    - Lambda Switch

Biología Sistémica

Hemos encontrado algunas carencias de ntcc para esta tarea

- Información cuantitativa explícita
  - Soporte para Ecuaciones Diferenciales
  - Inclusión de información probabilística (hay algunos avances)
  - Robustecer los sistemas de restricciones sobre números reales
- Mejoramiento de las herramientas de simulación existentes
- Integración de ayudas al usuario: editores (gráficos), soporte para simulaciones, etc.

- Esperamos modelar sistemas reales complejos
  - Refinar tanto los fundamentos teóricos como las herramientas existentes

Seguridad

- La expansión de las redes de comunicación trae consigo nuevos riesgos en la transmisión de la información
  - Internet, Aplicaciones Peer-to-Peer
- *Protocolos de seguridad*
  - Secuencias precisas de acciones que buscan asegurar un conjunto de propiedades de seguridad durante la transmisión de información sensible en ambientes inseguros
  - Algunas propiedades:
    - Anonimidad
    - Secrecidad
    - Autenticidad
    - No trazabilidad
- Idea: Modelar los participantes del protocolo como procesos concurrentes.

Seguridad

- Las pruebas ordinarias de la ingeniería de software no aseguran que los protocolos sean seguros.
  - Protocolo Needham-Schroeder.
- Verificar protocolos puede llegar a ser muy difícil:
  - La criptografía por si sola no es garantía de seguridad en los protocolos ---la lógica inherente a los protocolos debe garantizar su fiabilidad.
  - Los protocolos deben analizarse en ambientes concurrentes, en los cuales las capacidades de posibles atacantes sean modeladas de forma precisa.

Seguridad

- Suposiciones básicas en el análisis
  - Todo mensaje alguna vez transmitido puede ser recordado, es decir, los mensajes son *persistentes*
  - Los atacantes:
    - Pueden escuchar los mensajes que transitan en la red
    - Pueden recordar los mensajes transmitidos, modificarlos e incluir nuevos mensajes
    - Operativamente, tienen las mismas capacidades que un agente bien intencionado
  - Entre mayor es el poder asociado al atacante, más realistas serán las conclusiones obtenidas (Análisis peor caso)

Seguridad

Dos abstracciones clave

- Persistencia de los mensajes como un almacén de restricciones de seguridad
  - Indispensable: Inferencia sobre mensajes.

    Ej. Si el almacén contiene tanto un mensaje encriptado y la llave de encripción, puede deducirse el mensaje
- Protocolos ejecutándose infinitamente con operaciones “sofisticadas” de entrada
  - Idea: una operación que detecte la presencia de una pieza de información determinada en el almacén y reaccione ante ello

Seguridad

A partir de estas abstracciones hemos iniciado el diseño de **Secure CCP (SCCP)**

- Un lenguaje basado en CCP, específicamente orientado para modelar y verificar protocolos
- Actualmente tenemos
  - Definición de la sintaxis e intuiciones avanzadas sobre la definición de la semántica
  - Resultados teóricos positivos sobre el uso de lógica de primer orden para verificar procesos persistentes
  - Una herramienta prototipo para verificación de protocolos en SCCP, implementada en Mozart.

Seguridad

Retos inmediatos

- Especificación de los componentes semánticos (posiblemente abstractos) así como de los elementos de verificación, para el caso puntual de seguridad.
  - Ej. Qué significa, en el contexto de CCP, un *ataque de seguridad*?
  - Ej. Qué noción de *tiempo* es la más apropiada?
- Modelar protocolos reales que aprovechen las abstracciones básicas y permitan refinar la teoría
  - Protocolos para internet
- Implementación de una herramienta completa que permita la descripción, simulación y verificación de protocolos

Interacción Multimedial

Relacionando dos nociones de interacción

- Los avances en los nuevos medios han permitido el desarrollo de artefactos interactivos creativos
  - Narración interactiva, juegos por computador
- El reto es *inducir reacciones emotivas* y *comunicar contenidos* en estos artefactos, de forma similar a los elementos tradicionales de interacción

La interacción multimedia involucra sistemas que interactúan con el usuario de forma creativa

- Aprendiendo y manteniendo un modelo del estilo de comportamiento del usuario
- Utilizando dicho estilo para generar formas personalizadas de narración, presentación e interacción

Interacción Multimedial

Aunque la interacción multimedia involucra varias disciplinas (e.g., computación gráfica, generación de audio, procesamiento de señales) el *núcleo* de un sistema de interacción está compuesto de una arquitectura de *agentes concurrentes* que debe ser coherente y poderosa.

Sin embargo, poco se sabe sobre formalismos que hagan creativos a estos agentes

- generando formas creibles y sorprendentes de narración
- aprovechando un conocimiento adquirido en interacciones pasadas

Interacción Multimedial

- La interacción modelada por los cálculos de procesos puede aplicarse para modelar interacciones comunes en la multimedia
- Los cálculos basados en restricciones son buenos candidatos para formalizar esta relación:
  - Permiten modelar declarativamente, por medio de restricciones, condiciones complejas sobre las interacciones
  - El control de ntcc sobre el tiempo y el no-determinismo permite generar secuencias narrativas que involucren comportamiento innovador, que guarden consistencia con interacciones actuales y pasadas.

Interacción Multimedia

- Nuestras experiencias en este campo incluyen
  - Un improvisador musical virtual que aprende de un músico humano, modela su estilo y ejecuta improvisaciones novedosas que son consistentes con el modelo.

Sin embargo, hay algunos aspectos prácticos y teóricos que deben resolverse

- Sincronización de agentes ntcc en tiempo real
  - Polifonía de personajes virtuales
  - Múltiples puntos de vista, como agentes que acceden a contextos pasados
- Modelar la interrupción de procesos
- Incluir técnicas de aprendizaje (inteligencia artificial) en ntcc

Oportunidades

REACT supone una amplísima gama de posibilidades de trabajo, que pueden ajustarse según el interés

- Complejidad variable
  - proyectos dirigidos
  - tesis de pregrado
  - investigación avanzada, a nivel de postgrado (maestría o doctorado)
- Orientaciones diferentes:
  - Aspectos teóricos
    - Modelamiento y verificación de sistemas reales
    - Extensión y/o modificación de formalismos
  - Aspectos prácticos
    - Implementación de herramientas
    - Reingeniería de librerías
- Software o hardware

Varias combinaciones son posibles.

Más información

Grupo de Investigación AVISPA

URL: http://avispa.puj.edu.co

Director: Camilo Rueda

Lista de correo: avispa@googlegroups.com