

Sistemas Multiagente

Javier Béjar

ECSDI - 2021/2022 2Q

CS-GEI-FIB 



Introducción

Tendencias en la historia de la computación

- ⊙ Ubicuidad
- ⊙ Interconexión
- ⊙ Inteligencia
- ⊙ Delegación
- ⊙ Orientación a las personas

(“Introduction to Multi-Agent Systems”, M. Wooldridge, 2001)

⊙ Ubicuidad

- La capacidad de cómputo se ha incluido en multitud de elementos gracias a su abaratamiento (IoT)
- El aumento de su potencia permite que la sofisticación (e inteligencia) sea ubicua

⊙ Interconexión

- Los sistemas ya no son elementos aislados, están conectados en red formando sistemas distribuidos
- Esto lleva a la idea de modelar la computación como un proceso de interacción (**computation as interaction**)

⊙ Inteligencia

- La **complejidad de las tareas** automatizables ha ido creciendo (llegando a poder considerarse *inteligentes*)

⊙ Delegación

- Más tareas **se realizan automáticamente** (sin supervisión), incluso en tareas críticas

⊙ Orientación a las personas

- La creación de programas se basa en **abstracciones** y metáforas de cada vez más **alto nivel** (menos centrada en el computador, más cercana a nuestra visión)

Agente: primera definición

*Un sistema computacional capaz de **actuar** de manera **independiente** como **representante** de su usuario (**satisfaciendo** unos **objetivos** de diseño y **sin supervisión**)*

Sistema multiagente

Un sistema computacional compuesto de múltiples agentes que interaccionan entre ellos

- ⊙ Cada agente tendrá sus propios **objetivos** y **motivaciones**
- ⊙ El éxito de la interacción requerirá de **cooperación**, **coordinación** y **negociación** (precisamente las cualidades de las personas)

- ⊙ Esta metáfora nos lleva a plantear los sistemas software basados en agentes desde dos perspectivas: **Individuos** y **sociedades**
 1. **Diseño de agentes**: ¿Cómo diseñamos agentes capaces de resolver de manera autónoma las tareas que se les delegan?
 2. **Diseño de sociedades**: ¿Cómo diseñamos agentes capaces de interaccionar con otros de manera que resuelvan sus tareas, especialmente en caso de objetivos conflictivos?

Agentes Inteligentes

Agente: segunda definición

*Sistemas computacionales capaces de realizar **acciones** de manera **autónoma** en algún **entorno**, con el propósito de **alcanzar** una serie de **objetivos** que tiene **delegados***

- ⊙ El principal interés de los agentes es que son **autónomos** (capaces de actuar de manera independiente)
- ⊙ Un agente está fuertemente ligado y en continua interacción con su entorno:

percepción \mapsto decisión \mapsto acción \mapsto percepción \mapsto decisión \mapsto ...

- ⊙ Agentes Simples: Termostato, demon unix
- ⊙ No estamos interesados en agentes simples

Agente: tercera definición

Sistemas computacionales capaces de realizar **acciones** de manera **autónoma** y **flexible** en algún **entorno**, con el propósito de **alcanzar** una serie de **objetivos** que tiene **delegados**

- ⊙ Entendiendo como flexible: **Reactivo, proactivo y social**

- ⊙ Un sistema **reactivo** ha de mantener una interacción continua con el entorno y **responder a los cambios** que ocurren (a tiempo)
- ⊙ En un **entorno fijo**, un agente **no debe preocuparse** del resultado de sus acciones, puede actuar sin pensar en las consecuencias

- ⊙ Los entornos reales (los interesantes) son dinámicos, sus elementos cambian, su información es incompleta y/o incierta
- ⊙ En entornos dinámicos se ha de tener en cuenta la posibilidad de un resultado no esperado (preguntarse si debe realizarse la acción prevista)

- ⊙ Reaccionar al entorno es fácil
 - p.e., Estimulo \implies acción respuesta
- ⊙ Pero queremos que los agentes **hagan cosas por nosotros**.
- ⊙ Esto implica un **comportamiento dirigido por objetivos**.
- ⊙ **Proactividad** = Generar e intentar cumplir objetivos, no dirigidos únicamente por eventos, **tomando la iniciativa**
- ⊙ Esto implica el poder y saber reconocer oportunidades (cuándo se puede actuar, planificar)

- ⊙ Necesitamos que un agente **reaccione** apropiadamente a los **cambios** en el entorno (reactividad)
- ⊙ Necesitamos que un agente sea capaz de **cumplir objetivos** a largo plazo (proactividad)
- ⊙ Ambas características **pueden interferir**.
- ⊙ El conseguir una combinación adecuada de ambas es un problema no resuelto

- ⊙ El mundo real es un entorno **multi-agente**, no es posible obtener los objetivos propios sin **considerar** los de **otros**
- ⊙ Algunos objetivos solo se pueden cumplir con la **interacción** de otros
- ⊙ La **habilidad social** en agentes es la capacidad de interactuar con otros agentes (incluidos humanos) vía **cooperación**, **coordinación** y **negociación**
- ⊙ Esto implica el tener que usar algún **lenguaje de comunicación**

⊙ Objetos y agentes:

- Encapsulan un estado
- Se comunican por paso de mensajes
- Tienen métodos que corresponden con las acciones que se pueden realizar según su estado

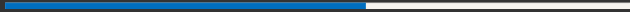
⊙ Pero los agentes son:

- Autónomos, deciden por si mismos si actúan o no
- Inteligentes, capaces de comportamientos flexibles
- Activos

“Los objetos lo hacen gratis, los agentes porque quieren y por dinero”

(M. Woolridge, Introd. to Multiagent Systems)

Arquitecturas abstractas de agentes



⊙ Arquitecturas reactivas puras

Los agentes poseen sensores y actuadores conectados al entorno. La conducta del agente se basa en estímulo-respuesta, la conexión entre sensores-actuadores hace **emerger** una conducta inteligente

⊙ Arquitecturas reactivas con estado interno

Los agentes tienen además un modelo del entorno que utilizan para decidir la reacción al estímulo.

- ⊙ **Arquitecturas deliberativas** (orientadas por objetivos)

Poseen una representación interna del mundo, siguen una aproximación simbólica y su funcionamiento se basa en el razonamiento (lógicas)

- ⊙ **Arquitecturas híbridas**

Diferentes capas de decisión combinan las dos aproximaciones. El objetivo es obtener las ventajas de cada una evitando los inconvenientes

- ⊙ Agentes que **deciden** sus acciones **independientemente de la historia**, sus acciones solo dependen del presente (no hay memoria)
- ⊙ Sus decisiones se basan en reglas simples que hacen coincidir las observaciones del estado con las decisiones:

$$\textit{Accion} : \textit{Entorno} \Rightarrow \textit{Actuacion}$$

- ⊙ Por ejemplo, un termostato

- ⊙ Agentes que tienen una **estructura interna** que representa información del **estado** y su historia
- ⊙ Poseen una función que percibe e interpreta el entorno:

$$\textit{Observar} : \textit{Entorno} \Rightarrow \textit{Percepcion}$$

- ⊙ El estado interno es usado para decidir la actuación:

$$\textit{Accion} : \textit{Estado} \Rightarrow \textit{Actuacion}$$

- ⊙ El estado es actualizado combinando la percepción y el estado interno actual:

$$\textit{Siguiente} : \textit{Estado} \times \textit{Percepcion} \Rightarrow \textit{Estado}$$

- ⊙ Los **agentes** se hacen para que puedan **resolver tareas por nosotros**
- ⊙ Esas tareas son **especificadas** por nosotros.
- ⊙ Queremos decirle al agente **qué hacer**, pero **no cómo hacerlo** (declarativo vs imperativo)
- ⊙ El agente debe tener capacidad para **elegir sus objetivos** y como conseguirlos

- ⊙ Basadas en una **visión simbólica de la IA**
- ⊙ Las decisiones se basan en la lógica simbólica
- ⊙ Se tiene un modelo del entorno y se actúa según ese conocimiento
- ⊙ **Problemas:**
 - Cómo representar del mundo exterior a partir de formalismos lógicos (**transductor problem**)
 - Cómo resolver el proceso de razonamiento que lleva a las decisiones (**representation/reasoning problem**)

- ⊙ Razonamiento basado el denominado **razonamiento práctico**
- ⊙ Modela el proceso que realizamos al decidir qué acción llevamos a cabo cada momento para perseguir unos fines
- ⊙ Basado en dos procesos
 - Decidir qué objetivos queremos conseguir (**Deliberación**)
 - Decidir cómo conseguirlos (**Razonamiento de medios fines**)
- ⊙ Los agentes se definen en función de: **Creencias**, **Deseos** e **Intenciones**

Sistemas Multiagente - Estándares

- ⊙ La interacción social entre agentes obliga a definir y desarrollar arquitecturas que soporten esta dimensión
- ⊙ Es necesaria una capa intermedia entre los agentes que permita la **interconexión/organización/comunicación**
- ⊙ Este software de soporte se denomina **plataformas de agentes**
- ⊙ Este middleware correspondería a la infraestructura definida por SOA RM/RAF

- ⊙ **FIPA** (Foundation for Intelligent Physical Agents) es un grupo de estandarización de IEEE que ha definido un conjunto de estándares sobre agentes
- ⊙ FIPA definió una **arquitectura abstracta** que debería seguir toda implementación de una plataforma multiagente
- ⊙ Está compuesta por:
 - Un directorio de agentes
 - Un directorio de servicios
 - Un mecanismo de transporte de mensajes
 - Un lenguaje de comunicación de agentes (ACL)

Temas/preguntas en sistemas
multiagente

- ⊙ Cómo diseñar y construir sistemas multiagente en la práctica
- ⊙ Cómo describir formalmente sistemas multiagente y la interacción entre agentes y como asegurar que están correctamente especificados
- ⊙ Cómo habilitar a los agentes para que descompongan sus tareas y objetivos (y asignar subtareas a otros agentes) y sintetizar resultados parciales
- ⊙ Cómo encontrar un compromiso entre coste computacional local y comunicación

- ⊙ Cómo habilitar a los agentes para representar y razonar sobre el estado y sus interacciones
- ⊙ Cómo habilitar a los agentes para representar y razonar sobre las acciones, planes y conocimiento de otros agentes para interaccionar con ellos
- ⊙ Cómo implementar en un sistema multiagente procesos inteligentes como resolución de problemas, planificación, toma de decisiones y aprendizaje

- ⊙ Cómo habilitar a los agentes para comunicarse, qué lenguajes y protocolos usar
- ⊙ Como formar y disolver estructuras organizativas para cumplir metas y objetivos específicos
- ⊙ Cómo habilitar a los agentes para reconocer y solucionar conflictos entre ellos
- ⊙ Cómo evitar o mitigar comportamientos indeseados en el sistema (caos)