



Desarrollo Basado en Agentes

Tema 2. Autonomía (II)

Luis Castillo Vidal Dpto Ciencias de la Computación e I.A. 2016-2017



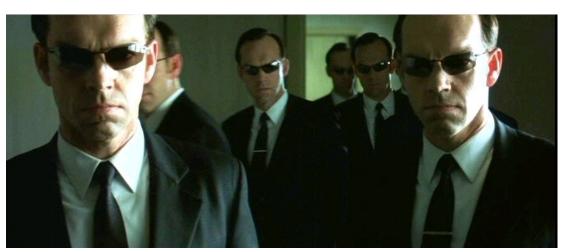
Esquema del tema

- Motivación
- Modelo básico de agentes que actúan
- Agentes deductivos
- Agentes de razonamiento práctico
- Agentes reactivos e híbridos



- Agentes de razonamiento simbólico
 - a. El agente contiene un modelo simbólico
 - i. De sí mismo (capacidades)
 - ii. Del entorno → ¿de otros agentes?
 - b. Las decisiones están dirigidas por procesos de razonamiento simbólico
 - c. <u>Comunmente</u> son modelos y razonamiento basados en lógica (predicados, proposicional)
- Independientes del lenguaje de programación

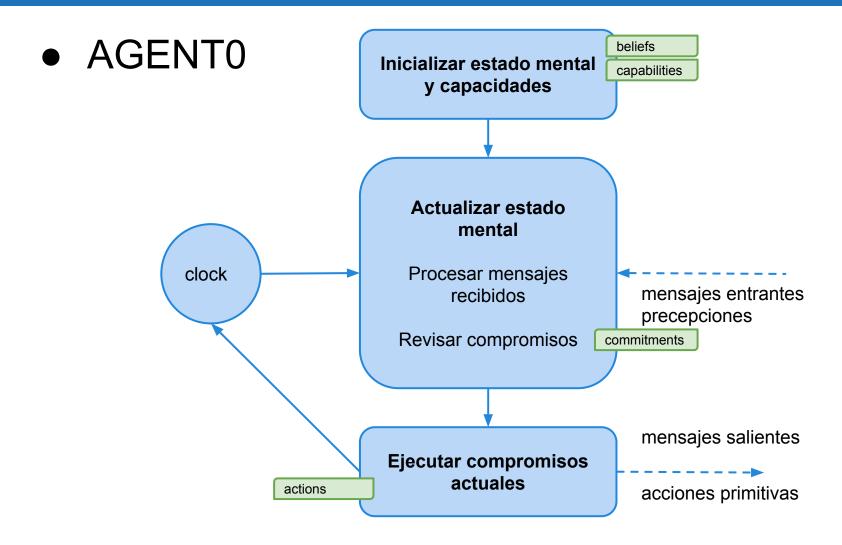
- "Teletransportables": si un agente es su KB, entonces ésta se puede transmitir (total o parcialmente) de un agente a otro
 - a. Agentes persistentes
 - b. Clonar agentes
 - c. Inyección de conocimiento



- Idea general
 - a. Sea KB = H ∪ R
 - i. $H (hechos) = \{h1, h2, ..., hn\}$
 - 1. hi = $literal(arg1, arg2, ..., argn) \rightarrow padre(X, Y)$
 - 2. La ontología
 - a. Las clases de objetos, sus propiedades y sus relaciones
 - 3. El estado de las cosas (Instanciaciones) → padre(Luis, Ana)
 - ii. R (reglas) = $\{r1, r2, ..., rn\}$
 - 1. Operadores lógicos ∧ V ¬ →
 - 2. Las acciones y capacidades del agente
 - b. El problema de actuar para conseguir unos objetivos se reduce a la demostración de teoremas
 - i. KB ⊢ objetivos
 - ii. mecanismos lógicos de inferencia
 - iii. La demostración → acción/plan-de-acción

Idea general





beliefs	commitments
capabilities	actions

?p → variable lógica libre para ser instanciada

Agentes deductivos

- AGENT0 (+ info <u>Shoham AOP</u>)
 - a. Estados mentales del agente
 - i. **Beliefs**: estado del mundo, *mental conditions* (MTC)
 - 1. (AND (OR (NOT ... (tiempo (predicado argumentos))
 - 2. (and (0 (abierta puerta1))(1 (posicion robot2 puerta1)))

```
Notación prefija LISP

3+4 ≡ (+ 3 4) abierta(puerta1) ≡ (abierta puerta1)
```

ii. Capabilities: Las acciones que puede realizar el agente y sus precondiciones

(CAN (DO tiempo accionprimitiva) mtc-precondition)

```
(CAN (DO 10:00 abrirpuerta) (and (10:00 (posicion myself ?p))
(10:00 (posicion puerta ?p)))
(CAN (DO ?anytime shutdown) true)
(CAN (DO ?t moveSE) (and (?t (bateria-nivel ?n))
(> ?n 1)))
```

beliefs	commitments
capabilities	actions

- a. Estados mentales del agente
 - i. **Commitments**: Reglas que expresan cuándo el agente se compromete a algo con otro (agente2), de motu propio o porque se lo pide otro agente (msg)
 - 1. (COMMIT msg mtc agente2 (DO tiempo accionprimitiva))

```
(COMMIT true (and (?t (bateria ?n)) (> ?n 10)) controler (DO ?t2 abrirpuerta))
```

- 2. Cada compromiso que se dispare modifica las creencias $(tiempo^c (CMT \ agente2 (DO \ tiempo^a \ aprim))) \rightarrow Beliefs$
- ii. Actions: lo que va a hacer el agente
 - 1. Acciones primitivas (DO tiempo accionprimitiva)
 - a. (DO 10:00 abrirpuerta)
 - 2. Acciones comunicativas
 - Acciones condicionadas

AGENT0



```
Resolución temporal s | m | h | d | y
```

```
(TIMEGRAIN m)
(CAPABILITIES
(CAN (DO ?t ON) true)
(CAN (DO ?t OFF) true))
(BELIEFS
(0 (T 17)) (0 (C 15))(0 (E OFF))
(0 (CAN (DO ?t ON) true)
(0 (CAN (DO ?t OFF) true)
(0 (CMT Timer) (DO 22:00 OFF))

)
(COMMITMENTS
```

Variables lógicas ?t

Introspección de las capabilities y commitments. Solo se muestan los hechos y los commitments por simplicidad

```
Agentes deductivos
                                                                                                                         Actualizar estado
                                                                                                                            mental
                                                                                                          clock
                                                                                                                         Procesar mensajes
                                                                                                                                        mensaies entrantes
                                                                                                                           recibidos
                                                                                                                                        precepciones
                                                                                                                        Revisar compromisos commitments
                                                                                                                                        mensajes salientes

    AGENTO

                                                                                                                       Eiecutar compromisos
                                                                                                                           actuales
                                                                                                                                        acciones primitivas
```

Agent Termostato v0 (TIMEGRAIN m) (CAPABILITIES (CAN (DO ?t ON) true) (CAN (DO ?t OFF) true)) (BELIEFS (0 (T 17)) (0 (C 15))(0 (E OFF)) (COMMITMENTS)

Revisar Compromisos en CLOCK = 0

Inicializar estado mental

y capacidades

capabilities

null

Ejecutar Compromisos para CLOCK = 0

null

```
Agentes deductivos
                                                                                                                         Actualizar estado
                                                                                                                            mental
                                                                                                          clock
                                                                                                                         Procesar mensajes
                                                                                                                                        mensaies entrantes
                                                                                                                           recibidos
                                                                                                                                        precepciones
                                                                                                                        Revisar compromisos commitments
                                                                                                                                        mensajes salientes

    AGENT0

                                                                                                                       Eiecutar compromisos
                                                                                                                           actuales
                                                                                                                                        acciones primitivas
```

```
Agent Termostato v0
(TIMEGRAIN m)
(CAPABILITIES
      (CAN (DO ?t ON) true)
      (CAN (DO ?t OFF) true))
(BELIEFS
      (1 (T 18)) (1 (C 15))(1 (E OFF))
(COMMITMENTS)
```

Revisar Compromisos en CLOCK = 1

Inicializar estado mental

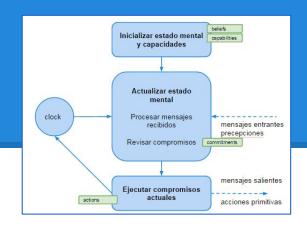
y capacidades

capabilities

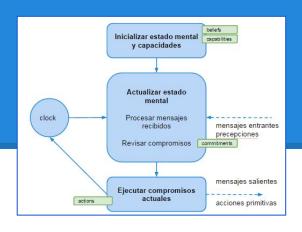
null

Ejecutar Compromisos para CLOCK = 1

null



```
Agent Termostato v1
(TIMEGRAIN m)
(CAPABILITIES
      (CAN (DO ?t ON) true)
      (CAN (DO ?t OFF) true))
(BELIEFS
      (0 (T 17)) (0 (C 15))(0 (E OFF))
(COMMITMENTS
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (DO ?t ON))
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostato (DO ?t OFF))
      (COMMIT true true Timer (DO 23:59 OFF))
                                    Consulta estado cognitivo
                                    (tiempo (pred arg))
                                                                   → Estado
                                                                   → Commitment
                                    (tiempo (CMT agente accion)
                                    (tiempo (CAN accion mtc))
                                                                   → Capability
```



```
Agent Termostato v1

(TIMEGRAIN m)

(CAPABILITIES

(CAN (DO ?t ON) true)

(CAN (DO ?t OFF) true))

(BELIEFS

(0 (T 17)) (0 (C 15))(0 (E OFF))

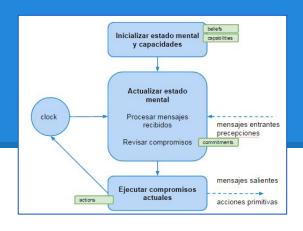
)

(COMMIT MENTS

(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (DO ?t ON))

(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostato (DO ?t OFF))

(COMMIT true true Timer (DO 23:59 OFF))
```



```
Revisar Compromisos en CLOCK = 0
Agent Termostato v1
                                                                         (0 (CMT Termostato (DO 0 ON)))
(TIMEGRAIN m)
                                                                         (0 (CMT Timer (DO 23:59 OFF)))
(CAPABILITIES
      (CAN (DO ?t ON) true)
                                                                         → BELIEFS
      (CAN (DO ?t OFF) true))
(BELIEFS
      (0 (T 17)) (0 (C 15))(0 (E OFF)) (0 (CMT Termostato (DO 0 ON)))
      (0 (CMT Timer (DO 23:59 OFF)))
(COMMITMENTS
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (DO ?t ON))
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostato (DO ?t OFF))
      (COMMIT true true Timer (DO 23:59 OFF))
```

```
Inicializar estado mental y capacidades

Actualizar estado mental Procesar mensajes entrantes precepciones Revisar compromisos commitments

Ejecutar compromisos actuales

Ejecutar compromisos acciones primitivas
```

```
Agent Termostato v1

(TIMEGRAIN m)
(CAPABILITIES
(CAN (DO ?t ON) true)
(CAN (DO ?t OFF) true))
(BELIEFS
(0 (T 17)) (0 (C 15))(0 (E OFF)) (0 (CMT Termostato (DO 0 ON)))
(0 (CMT Timer (DO 23:59 OFF)))
)
(COMMITMENTS
(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostato (COMMIT true true Timer (DO 23:59 OFF))

(0 (CMT termostato (DO 0 ON))))
```

```
Inicializar estado mental y capacidades

Actualizar estado mental

Procesar mensajes recibidos mental

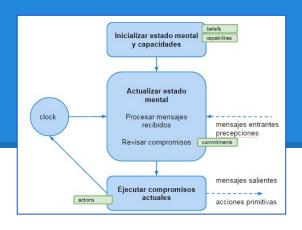
Procesar mensajes entrantes precepciones

Revisar compromisos commitments

Ejecutar compromisos actuales

acciones primitivas
```

```
Agent Termostato v1
(TIMEGRAIN m)
(CAPABILITIES
      (CAN (DO ?t ON) true)
      (CAN (DO ?t OFF) true))
(BELIEFS
      (0 (T 17)) (0 (C 15))(0 (E OFF))
      (0 (CMT Timer (DO 23:59 OFF)))
(COMMITMENTS
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (D
                                                                           Ejecutar Compromisos para CLOCK = 0
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostato
      (COMMIT true true Timer (DO 23:59 OFF))
                                                                           (0 (CMT termostato (DO 0 ON)))
                                                                           El termostato se activa
                                                                           Aún no hemos actualizado las percepciones
```



```
Agent Termostato v1

(TIMEGRAIN m)
(CAPABILITIES
(CAN (DO ?t ON) true)
(CAN (DO ?t OFF) true))

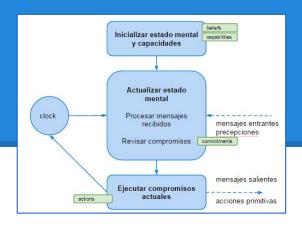
(BELIEFS

(1 (T 16)) (1 (C 15))(1 (E ON))
(0 (CMT Timer (DO 23:59 OFF)))

)

(COMMITMENTS
(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (DO ?t ON))
(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostato (DO ?t OFF))
(COMMIT true true Timer (DO 23:59 OFF))

)
```



```
Agent Termostato v1

(TIMEGRAIN m)
(CAPABILITIES
(CAN (DO ?t ON) true)
(CAN (DO ?t OFF) true))

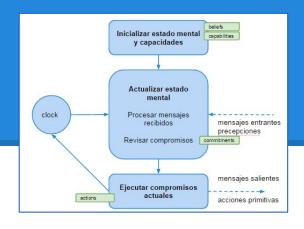
(BELIEFS

(1 (T 16)) (1 (C 15))(1 (E ON))(1 (CMT Termostato (DO 1 ON)))
(0 (CMT Timer (DO 23:59 OFF)))

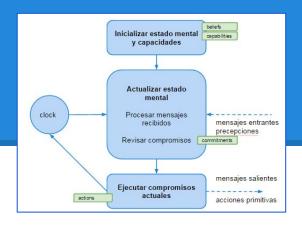
)

(COMMITMENTS
(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (DO ?t ON))
(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostato (DO ?t OFF))
(COMMIT true true Timer (DO 23:59 OFF))

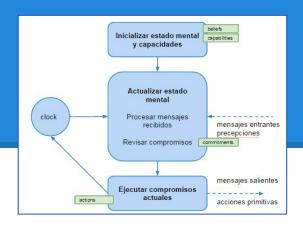
)
```



```
Agent Termostato v1
(TIMEGRAIN m)
(CAPABILITIES
      (CAN (DO ?t ON) true)
      (CAN (DO ?t OFF) true))
(BELIEFS
      (1 (T 16)) (1 (C 15))(1 (E ON))(1 (CMT Termostato (DO 1 ON)))
      (0 (CMT Timer (DO 23:59 OFF)))
(COMMITMENTS
                                                                          Ejecutar Compromisos para CLOCK = 1
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostat
                                                                          (1 (CMT termostato (DO 1 ON)))
      (COMMIT true true Timer (DO 23:59 OFF))
```



```
Agent Termostato v1
(TIMEGRAIN m)
(CAPABILITIES
      (CAN (DO ?t ON) true)
      (CAN (DO ?t OFF) true))
(BELIEFS
      (1 (T 16)) (1 (C 15))(1 (E ON))
      (0 (CMT Timer (DO 23:59 OFF)))
(COMMITMENTS
                                                                           Ejecutar Compromisos para CLOCK = 1
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostat
                                                                           (1 (CMT termostato (DO 1 ON)))
      (COMMIT true true Timer (DO 23:59 OFF))
                                                                           El termostato se activa
                                                                           Aún no hemos actualizado las percepciones
```



```
Agent Termostato v1

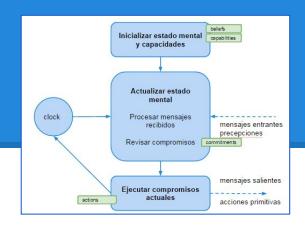
(TIMEGRAIN m)
(CAPABILITIES
(CAN (DO ?t ON) true)
(CAN (DO ?t OFF) true))

(BELIEFS
(2 (T 15)) (2 (C 15))(2 (E ON))
(0 (CMT Timer (DO 23:59 OFF)))

)

(COMMITMENTS
(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (DO ?t ON))
(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostato (DO ?t OFF))
(COMMIT true true Timer (DO 23:59 OFF))

)
```



```
Agent Termostato v1

(TIMEGRAIN m)
(CAPABILITIES
(CAN (DO ?t ON) true)
(CAN (DO ?t OFF) true))

(BELIEFS
(2 (T 15)) (2 (C 15))(2 (E ON))(2 (CMT Termostato (DO 2 OFF)))
(0 (CMT Timer (DO 23:59 OFF)))

)

(COMMITMENTS
(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (DO ?t ON))
(COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostato (DO ?t OFF))
(COMMIT true true Timer (DO 23:59 OFF))

)
```

```
Inicializar estado mental y capacidades

Actualizar estado mental

Procesar mensajes recibidos mental

Procesar mensajes entrantes precepciones

Revisar compromisos commitments

Ejecutar compromisos actuales

acciones primitivas
```

```
Inicializar estado mental y capacidades

Actualizar estado mental

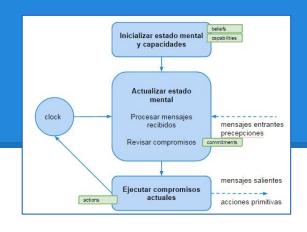
Procesar mensajes recibidos mensajes entrantes precepciones commitments

Revisar compromisos commitments

Ejecutar compromisos actuales

acciones primitivas
```

```
Agent Termostato v1
(TIMEGRAIN m)
(CAPABILITIES
      (CAN (DO ?t ON) true)
      (CAN (DO ?t OFF) true))
(BELIEFS
      (2 (T 15)) (2 (C 15))(2 (E ON))
      (0 (CMT Timer (DO 23:59 OFF)))
(COMMITMENTS
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (> ?TE ?CO)) Termostato (D
                                                                           Ejecutar Compromisos para CLOCK = 2
      (COMMIT true (and (?t (T ?TE)) (?t (C ?CO)) (<= ?TE ?CO)) Termostato
      (COMMIT true true Timer (DO 23:59 OFF))
                                                                           (2 (CMT termostato (DO 2 OFF)))
                                                                           El termostato se desactiva
                                                                           Aún no hemos actualizado las percepciones
```



AGENTO

a. Acciones primitivas

(DO tiempo accionprimitiva)

- b. Acciones comunicativas
 - i. (REQUEST tiempo agente2 accion)



- ii. (INFORM tiempo agente2 mtc)
- c. Acciones condicionadas
 - i. (IF mtc <u>accion</u>)
- d. Condiciones sobre los mensajes recibidos (COMMIT *msg mtc agente2* (DO *tiempo accionprimitiva*))
 - i. (agentesender REQUEST accion)
 - ii. (agentesender INFORM mtc)

AGENTO

a. Agentes "Termostato", "Timer", "User"

```
Agent Termostato v3
(TIMEGRAIN m)
(CAPABILITIES
      (CAN (DO ?t ON) true)
      (CAN (DO ?t OFF) true))
      (CAN (REQUEST ?t ?X ?A) true)
      (CAN (INFORM ?t ?X ?mtc) true)
(BELIEFS
      (3 (T 15)) (3 (C 15))(3 (E OFF))
(COMMITMENTS
      (COMMIT true (and (B (and (?t (T ?TE))) (?t (C ?CO)))) (> ?TE ?CO)) Termostato (DO ?t ON))
      (COMMIT true (and (B (and (?t (T ?TE))) (?t (C ?CO)))) (<= ?TE ?CO)) Termostato (DO ?t OFF))
      (COMMIT (Timer REQUEST (DO ?t OFF)) true Timer (DO ?t OFF))
      (COMMIT (?anyagent REQUEST (INFORM ?t ?anyagent ?mtc)) true ?anyagent (INFORM ?t ?anyagent ?mtc))
      (COMMIT true (and (?t (E OFF)) (?t (T ?TE))) (?t (C ?CO)) (> ?TE ?CO)) User (INFORM ?t User (?t (E BAD)))
```

- AGENT0: Buenas propiedades
 - Consistencia de las creencias (beliefs)
 - No debe haber hechos inconsistentes
 - (BELIEFS (0 (not (abierta puerta))) (0 (abierta puerta)))
 - Y si las hay ¿cuál eliminar?
 - Actuación de buena fé
 - El agente no puede comprometerse a hacer algo que no está entre sus capacidades
 - El agente no puede informar sobre la veracidad de algo que desconoce

 El agente no puede comprometerse a hacer algo que no está entre sus capacidades

(COMMIT (?Ag REQUEST (DO ?t ?action)) (and [0 (CAN (DO ?t ?action) ?mtc] ?mtc) ?Ag (DO ?t ?action))

 El agente no puede informar sobre la veracidad de algo que desconoce

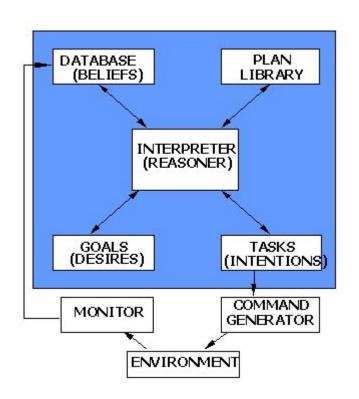
(COMMIT (?A REQUEST (INFORM ?t ?A2 ?mtc)) (and ?mtc (0 (CAN (INFORM ?AT ?AA ?AC)))) ?A (INFORM ?t ?A2 ?mtc))

- AGENT0: Buenas propiedades
 - Persistencia del estado mental
 - **Beliefs**: actualizar y eliminar aquellos hechos que dejan de ser ciertos
 - Capabilities: inicialmente son estáticas, pero en agentes avanzados podrían variar (aprender nuevas capacidades o eliminar aquellas obsoletas).
 - <u>"Ya sé Kung-Fu"</u> → ¡no tan <u>futurista!</u>
 - Commitments: persisten hasta que son completadas.
 - En algunos modelos de agentes, el agente que hace un REQUEST puede hacer un UNREQUEST cuando ya lo no necesita, liberando así al agente que se comprometió

redes

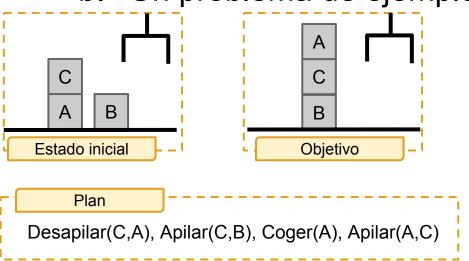
- Agentes deliberativos: Razonan sobre <u>secuencias de acciones</u> y su efecto en el entorno
- Asociada a agentes (intención)
- SIEMPRE sobre un modelo deductivo
- Entradas
 - i. Objetivos
 - ii. Estado actual → (i+ii) Problema
 - iii. Acciones → (iii) Dominio
- Salidas
 - i. Plan

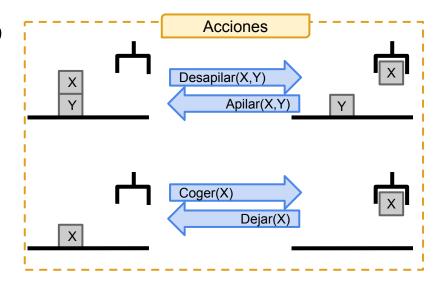
- Agentes BDI (+info <u>Wikipedia</u>)
 - Beliefs
 - Desires
 - Intentions



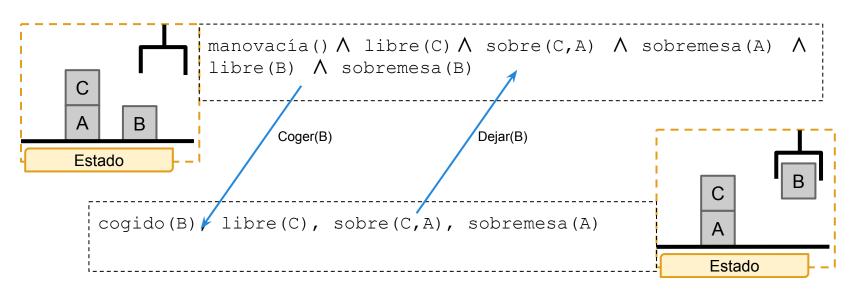
- Agentes que planifican

 BDI
 - a. El mundo de bloques
 - i. Uno (o más robots)
 - ii. Una serie de bloques
 - iii. Una superficie
 - b. Un problema de ejemplo

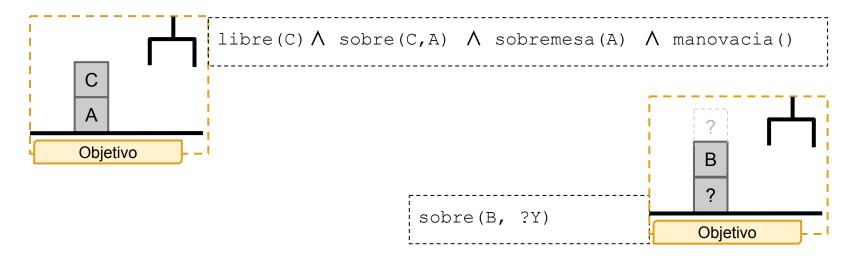




- Agentes que planifican
 - a. Lenguaje de representación
 - i. **De estados**. Suele estar basado en la lógica de predicados.
 - 1. Completitud → **Hipótesis de mundo cerrado** (CWA)
 - 2. En problemas con percepción real CWA ¡no existe!

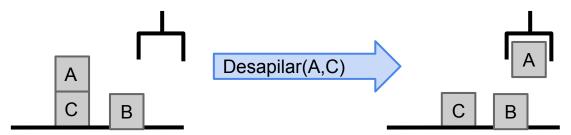


- Agentes que planifican
 - a. Lenguaje de representación
 - i. De objetivos. Suele estar basado en la lógica de predicados. Puede contener variables o estar parcialmente definido (conjunto de estados)

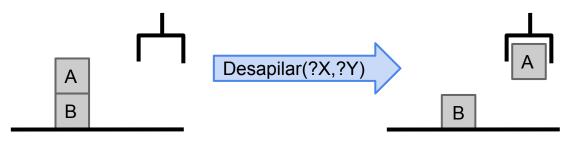


- Agentes que planifican
 - a. Lenguaje de representación
 - i. **De acciones**. Suele estar basado en la lógica de predicados.
 - Calculo situacional (lógica predicados)

Problema del marco → Intratabilidad



- Agentes que planifican
 - a. Lenguaje de representación
 - i. **De acciones**. Suele estar basado en la lógica de predicados.
 - Planificador STRIPS (basado en la lógica)
 - Listas de precondición, supresión y adición → Probl. del marco
 - Con variables (lógica de predicados)
 - Sin variables (lógica proposicional)



Acción Desapilar(?X,?Y)

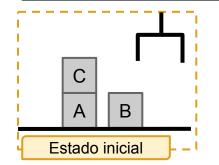
P: manovacia(), libre(?X), sobre(?X,?Y)

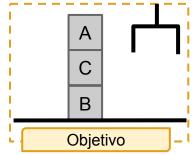
S: manovacia(), libre(?X), sobre(?X, ?Y)

A: cogido(?X), libre(?Y)

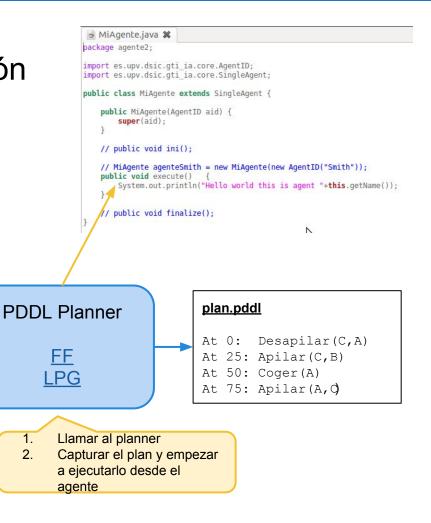
PDDL (Planning Domain Description Language)

```
Acciones posibles
domain.pddl
 ;; Ontología
(:types Block - Object)
(:predicates (libre ?X - Block) ... )
(:functions (peso ?X - Block) ... )
;; Capacidades
(:action Desapilar
 :parameters (?X ?Y - Block)
 :condition (and (manovacia) (libre ?X)
            (sobre ?X ?Y))
 :duration (= ?duration 25)
 :effects (and (not (manovacia))
                (not (libre ?X))
                (not (sobre ?X ?Y))
                (cogido ?X) (libre ?Y)))
```



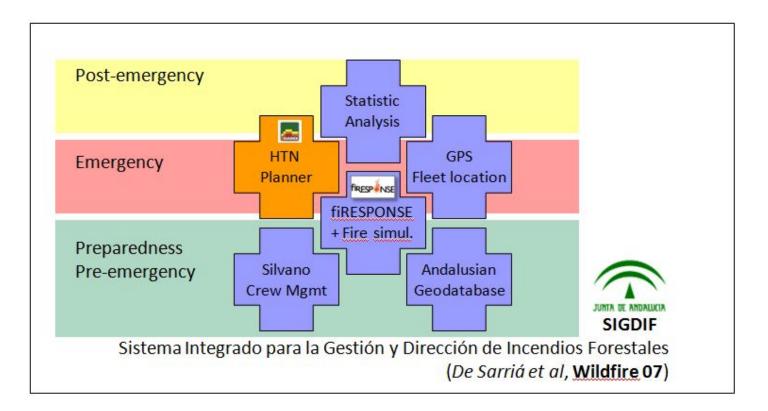


- Planificación
 - a. Lenguaje de representacióni. **De planes**.
 - b. Algoritmos

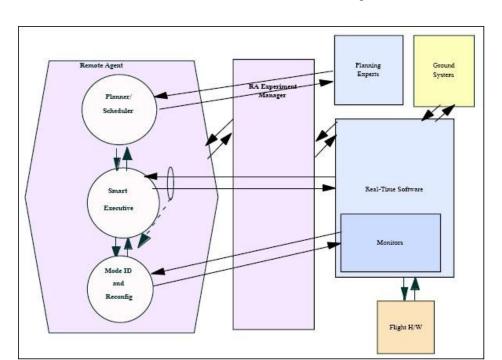




- lActive Decisor (planificación jerárquica)
 - Agente Reactivo/Deliberativo con aprendizaje
 - Silent Mode



- NASA Remote Agent Experiment (RAX'99)
 - Misión "Deep Space 1"
 - Un agente planificador que controló durante 3 días una sonda espacial
 - Recibió el conocimiento ya en órbita





- Inherentemente compleja. Factor
 - diferenciador de la inteligencia humana (y algunos primates)
 - Predicción del futuro
 - Uso de herramientas
 - Definición de procesos



- Computacional. No es un problema sencillo, requiere tiempo (<u>exponencial</u>)
 - o El mundo puede cambiar mientras responde
- Interferencias. El plan de un agente no debe interferir con otros planes (de él o de otros) → Proyecto comorbilidad
- Goal-oriented. Un agente no abandona sus objetivos si un plan falla → prueba otro plan para los mismos objetivos → ¿se puede abandonar un objetivo?

- Razonamiento sobre objetivos. Un agente siempre cree que sus objetivos son alcanzables
- Certeza. Un agente nunca intenta un plan si sabe que no funcionará
 - InfoSiadex → cambios de turno en retenes
- Eficacia. Si un agente tiene un plan es porque cree que tendrá éxito con el plan.

- Efectos laterales. Un agente no planifica los efectos laterales de un plan, pero algunos efectos laterales pueden ser beneficiosos para un plan.
- Completitud. Se espera que el estado inicial sea completo

Algunas ventajas

- Conocimiento transferible. De un agente a otro "sin más" que transferir la parte correspondiente de PDDL
- Real world. Algunos lenguajes son usados en aplicaciones reales muy poderosas → capacidad práctica del agente muy potente