

SISTEMA AFECTIVO PARA LA REGULACIÓN DE COMPORTAMIENTOS DE AGENTES ROBOTICOS

Carlos Domínguez, Houcine Hassan, Alfons Crespo
Departamento de Informática de Sistemas y Computadores (DISCA)
Universidad Politécnica de Valencia (UPV)
Camino de Vera, 14 - 46022 Valencia
{carlosd,husein,alfons}@disca.upv.es

Resumen

La arquitectura RTEA – Real Time Emotional Agent, define un marco para la especificación y ejecución de procesos de control en tiempo real, en el que explícitamente se modela la dinámica del problema y la dedicación necesaria de procesador.

RTEA intenta facilitar el desarrollo de aplicaciones de control que satisfagan los requerimientos críticos (seguridad e integridad) y optimicen la satisfacción de los requerimientos funcionales.

El sistema emocional distribuido, que define emociones “locales” a cada uno de los procesos activos, permite un modelado más sencillo de la organización mental. Esto presupone que existen reglas de interacción sencillas entre los distintos procesos, que una vez definidas y afinadas, permitirán una satisfacción promedio aceptable en el conjunto de requerimientos funcionales.

En este trabajo se plantean distintos tipos de interacción entre procesos simultáneamente ejecutados por el agente, así como distintos tipos de caracteres y rasgos (tipos de respuesta emocional) ante dichos conjuntos de procesos.

Palabras Clave: Emoción, Carácter, Motivación, Tiempo Real, Robótica.

1 INTRODUCCIÓN Y MOTIVACIÓN

Según recientes estudios [6], se prevé que en los distintos campos de aplicación de la robótica: producción avanzada, edificios inteligentes y sirvientes adaptativos, robots de servicio y de campo, robots humanoides, vehículos inteligentes y logística, etc., y a corto/medio plazo, se construyan robots cada vez más inteligentes y con mayores habilidades perceptivas y cognitivas para facilitar la interacción hombre-robot.

Los sistemas emocionales están siendo actualmente considerados como potentes modelos computa-

cionales para diseñar mejores robots y más inteligentes [1]. Se trata de una rama de la robótica que no ha hecho más que emerger, y con ella aparecen diversas cuestiones a estudiar relativas al diseño de agentes emocionales [3].

En el contexto de la neurociencia una emoción es la reacción organizada a un evento que es relevante para las necesidades, intereses o supervivencia del organismo [8]. Desde el punto de vista de la robótica una emoción podría verse como un subconjunto de motivaciones que se utilizan para modular dinámicamente el control del comportamiento manifestado por un agente robótico [3].

1.1 CARACTERÍSTICAS DEL AGENTE RTEA

Cuando un agente RTEA organiza la ejecución de sus procesos: (1) satisface requerimientos intrínsecos de seguridad y (2) maximiza la satisfacción de requerimientos funcionales.

Los requerimientos intrínsecos de seguridad se establecen especificando Pensamientos Responsables, que sólo tendrán lugar si previamente se determina que se tiene capacidad para llevarlos a término sin poner en compromiso la seguridad del agente o la de alguna parte protegida de su entorno.

Los requerimientos funcionales se establecen formulando Deseos. La formulación del deseo es flexible para que pueda adaptarse a la capacidad real del agente. La flexibilidad en la formulación se basa en un modelo de satisfacción (aceptabilidad de las soluciones), que permite negociar la dedicación de procesador para maximizar la satisfacción sin poner en peligro la satisfacción de los requerimientos intrínsecos de seguridad.

Cada Deseo provoca un Pensamiento que intenta satisfacerlo. El Pensamiento, como abstracción consciente de un proceso en ejecución, define explícitamente tres modelos: (1) un Modelo de Dinámica (MoDin), (2) un Modelo de Dedicación (MoDed) y (3) un Modelo de Satisfacción (MoSat)

Estos modelos permiten al agente adaptar sus deseos a sus capacidades (visión realista de su influencia en el entorno). Son modelos reversibles, de forma que pueden ser interrogados tanto para el Diseño como para la Previsión.

El MoDin define cómo las variables de interés en el dominio del problema (vi) cambian en el tiempo ($vi_0 \rightarrow vi_1$) en función de unas condiciones constantes (ck) también definidas en el dominio del problema.

El MoDed define la dedicación mental (capacidad de procesador asignada) necesaria para mantener las condiciones (ck) sin violar los requerimientos intrínsecos de seguridad.

El MoSat define el grado de satisfacción del deseo dada una situación (vi), real (actual) o imaginaria (previsión).

El Sistema de Atención negocia con los pensamientos aplicando un Modelo de Atención (MoAtenc) que básicamente establece una satisfacción mínima esperada para el deseo que cada pensamiento está persiguiendo en función de la motivación del pensamiento.

Las motivaciones de los pensamientos son establecidas por el sistema emocional. La Emoción se entiende como una alteración del ánimo (motivación de los pensamientos).

Las emociones intrínsecas a cada pensamiento establecen una motivación base, derivada de la importancia del deseo a satisfacer, y una modulación sobre dicha motivación base, que aplica el carácter del agente y una valoración de la ejecución del pensamiento y las expectativas de satisfacción del deseo que persigue. Así, valoraciones de urgencia o de oportunidad, pueden alterar la motivación base del pensamiento (importancia del deseo) y ‘aparentar’ que el deseo es más o menos importante.

Este tipo de ‘distorsiones’ son comunes en el comportamiento de los agentes biológicos reales, y dan como resultado ventajas o desventajas a cada agente dependiendo del tipo de entorno.

Son este tipo de distorsiones emocionales las que nos interesa estudiar en este trabajo. Así, se definirán los caracteres y rasgos del agente y mostrará la intervención de la actitud del agente en el proceso de regulación de su estado afectivo para el cumplimiento de sus tareas.

Después de la Introducción, en el apartado 2 se comenta los trabajos relacionados con esta temática. En el apartado 3 se realiza la descripción de la arquitectura RTEA. El apartado 4 detalla la

estructura del carácter del agente. Finalmente, las conclusiones se resumen en el apartado 5.

1.2 APLICACIONES PARA AGENTES RTEA

El Agente Asistente es una aplicación típica para la arquitectura RTEA. Se trata de una aplicación genérica que puede concretarse en una amplia gama de entornos prácticos: industriales, comerciales, domóticos, asistencia personal, etc.

Un agente asistente se desenvuelve en un entorno limitado y semi-estructurado, en el que emergen deseos y se activan pensamientos de satisfacción de los mismos. Los conjuntos de deseos y pensamientos están preestablecidos, pero son definidos de forma suficientemente general para que puedan concretarse de variadas formas.

Los deseos pueden emerger como consecuencia de órdenes directas que llegan a través de una interfaz con un operador, por ejemplo la orden “Ayúdame a colocar esto allí”, con lo que el agente formula un deseo de tipo “<objeto> en <lugar> colaborando_con <operario>”.

También pueden emerger como una concreción de deseos formulados de forma general, por ejemplo la orden general “Limpia cada derrame de aceite que encuentres en el taller” puede formularse en un deseo permanente “ \forall <lugar> limpio_de <aceite>” que se concreta en deseos concretos ante la detección de manchas de aceite reales “<lugar> limpio_de <aceite>”.

Finalmente los deseos pueden ser formulados de forma operativa por el propio agente. Si limpiar una mancha de aceite implica estar a una distancia práctica de la misma, los propios procesos mentales relacionados con limpiar mancha deben formular deseos del tipo “estar_en <lugar>”

El entorno dinámico obliga al agente asistente a actuar en tiempo real satisfaciendo requerimientos de seguridad. El conjunto de deseos simultáneos puede ser grande, con lo que el agente asistente debe determinar qué deseos satisfacer y con qué grado de satisfacción, limitándose a sus capacidades y cumpliendo los requerimientos de seguridad.

Para este tipo de aplicaciones se propone la arquitectura RTEA.

2 ANTECEDENTES

En la literatura aparecen varios modelos basados en emociones, pero la mayoría se centran en la interacción hombre-robot. Ronald Arkin propone un

modelo etológico basado en la mantis religiosa donde se controlan 3 variables motivacionales; hambre, miedo y deseo sexual asociadas a los diferentes comportamientos del robot [1].

Sony en colaboración con el equipo de Arkin diseñan el robot de entretenimiento Aibo [2]. Este se basa en 6 estado emocionales: felicidad, angustia, tristeza, miedo sorpresa, disgusto. Los comportamientos de Aibo se seleccionan en base a su estado motivacional.

La directora del departamento de "Social Robots" del MIT Media Lab., Cynthia Breazeal, desarrolla uno de los primeros robots sociales completos: Kismet [4]. Este incorpora un sistema emocional que regula el sistema motivacional mediante 3 dimensiones (excitamiento, valencia, actitud). El estado motivacional afecta a la interacción con los humanos mediante la activación de diferentes expresiones faciales.

Moshkina propone el modelo Tame para mejorar la interacción hombre-robot [7]. Tame dispone de 4 variables afectivas variables en el tiempo: actitud, rasgo, humor, emoción. El modelo obtiene información perceptual y modifica los parámetros de los comportamientos de interacción.

En general, los modelos descritos se limitan a mejorar la interacción hombre-robot como bien se puntualiza en [3, 4]. Estos modelos son suficientes para sistemas no móviles (p.e. Kismet) o que tengan un dinámica donde las restricciones de tiempo real no sean críticas (p.e Aibo). Sin embargo, un modelo de agente emocional en robótica móvil debe adaptarse en tiempo real a las condiciones del entorno en el que se mueve, basándose en sus capacidades físicas (actuadores) y mentales (procesos) [5].

3 ARQUITECTURA

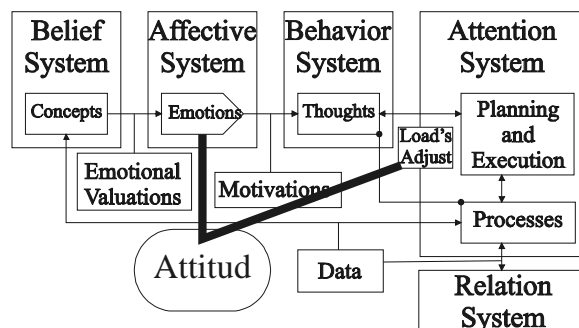


Figura 1: Subsistemas de la arquitectura RTEA

Un agente RTEA se compone de 5 subsistemas: Creencia, Afectivo, Comportamiento, Atención y Relación.

La Figura 1 muestra el flujo principal de información entre dichos subsistemas.

3.1 CONCEPTOS

El *Sistema de Creencias* mantiene una imagen lógica del entorno en forma de una estructura de datos. Los procesos en ejecución interrogan y actualizan esa imagen permanentemente.

Los elementos fundamentales del sistema de creencias son los Conceptos, que son abstracciones conscientes de los datos.

La conciencia acerca los datos se define mediante un conjunto de Valoraciones Conscientes de los mismos. Estas valoraciones son números reales en el rango [0,1] y pueden ser función tanto del valor actual del dato, como de la forma en que dicho dato se generó, o la forma en que dicho dato se espera utilizar.

3.1.1 Valoración de Confianza

Un ejemplo de valoración consciente muy importante es la Confianza. Ésta se utiliza como una medida de la aplicabilidad del dato. Una valoración de confianza igual a 1 indica que el dato es muy confiable y se puede aplicar con seguridad. Una valoración de confianza igual a 0 indica que el dato no es nada confiable y no se debe aplicar.

Los procesos que actualizan el valor de un dato en la imagen del entorno deben establecer la valoración de la confianza para el concepto consciente de dicho dato. Una valoración de confianza no es un simple valor que se asigna en un momento dado, sino una función, que incorpora el modelo de evaluación. De esta manera, cada vez que se interroga al concepto sobre su confianza, el valor se evalúa nuevamente.

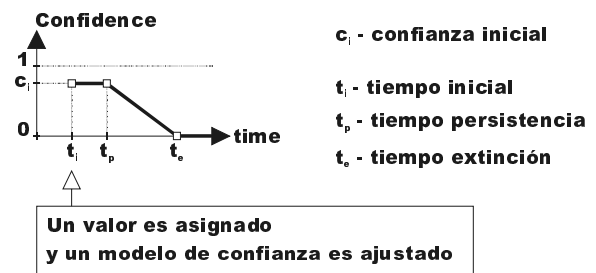


Figura 2: Modelo de evaluación de confianza

Un modelo sencillo de evaluación de la confianza puede ser una función sólo del tiempo. Esta función puede definir un modelo de degradación de la confianza con mayor o menor rapidez en función de las condiciones supuestas en el momento de la generación del valor como de las condiciones

previstas en un futuro hasta un horizonte dado. Realmente los modelos de evaluación deben ser lo más sencillos posibles, para minimizar el coste de evaluación. La Figura 2 muestra un modelo sencillo de evaluación de la confianza en forma de rampa retardada, en el que los parámetros son los tiempos de inicio (cuándo se define el modelo), de persistencia (hasta cuando se mantiene la confianza inicialmente otorgada), y de extinción (cuando la confianza es nula).

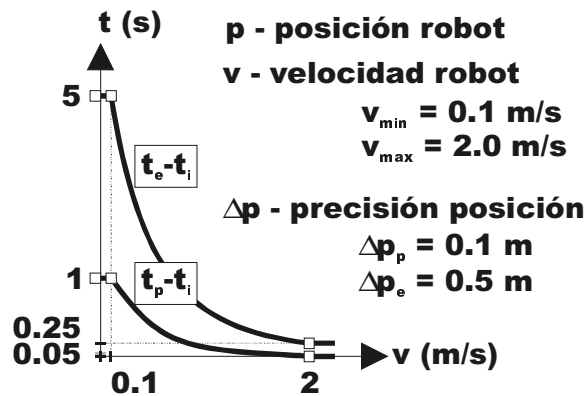


Figura 3: Definición del modelo de confianza

Por ejemplo, la confianza en la posición ocupada por un robot móvil podría evaluarse en función de la velocidad del robot en el momento en que se evaluó. Una velocidad alta propondría un modelo de degradación de la confianza muy rápido. El ajuste del modelo podría estar impuesto por requerimientos de precisión.

En la Figura 3 muestra un ejemplo de evaluación de los tiempos de persistencia y los tiempos de extinción para la definición de un modelo de evaluación de la confianza en la posición del robot.

3.1.2 Aplicación de las valoraciones conscientes

La valoraciones conscientes de los datos son simples medidas, que no implican indefectiblemente una influencia en el comportamiento del agente.

Estas valoraciones sí podrán sin embargo contribuir a la definición del estado emocional del agente y finalmente influir en su comportamiento si se definen y activan emociones basándose en ellas.

3.1.3 Objetividad y Subjetividad

Para poner claramente de manifiesto la influencia que el carácter del agente tiene en su comportamiento ante las situaciones en el entorno, quisiéramos considerar el grado de objetividad o subjetividad de las valoraciones.

De alguna manera nos gustaría decir que las valoraciones conscientes tendrían un apreciable grado de objetividad. Objetividad en el sentido de que dada una especificación de problema de aplicación, distintos ingenieros de programación de sistemas definirían modelos de evaluación consciente muy parecidos.

Pero estas valoraciones conscientes podrían contribuir de formas sensiblemente distintas al estado emocional y finalmente al comportamiento del agente, en función de la valoración emocional que se realice de las mismas.

Quisiéramos decir también que hay un grado de subjetividad en las valoraciones emocionales. El grado de subjetividad se manifestaría a través de la actitud del agente, derivada ésta de su carácter y de su estado de humor actual.

Un ingeniero de programación puede entonces definir distintos caracteres y estados de humor y esperar distintos resultados en el comportamiento del agente ante las mismas situaciones.

3.2 EMOCIONES

El *Sistema Afectivo* es el motor en la organización mental, y mantiene un conjunto de Emociones como mecanismo básico de alteración del ánimo.

El ánimo se define como el grado de Motivación de cada uno de los pensamientos activos.

Las emociones definen una condición, un estado y una respuesta emocionales.

Las condición para la activación de un estado emocional se basa en las contribuciones que ciertas valoraciones conscientes de conceptos puedan tener sobre dicho estado emocional, y la respuesta es el cambio en la motivación de los pensamientos.

El carácter y el humor del agente modulan la sensibilidad de las contribuciones y el sentido e intensidad de las respuestas emocionales.

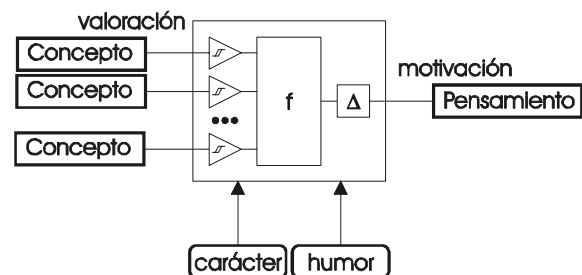


Figura 4: Emoción

El carácter del agente define la modulación emocional constante, mientras que el estado de humor define la modulación emocional variable.

Más adelante, el apartado 4 discute el mecanismo modulador del carácter en el sistema afectivo.

3.3 PENSAMIENTOS

El *Sistema de Comportamiento* define la conducta que el agente puede mostrar. Los comportamientos son procesos. La representación consciente de éstos son los Pensamientos.

La conciencia se materializa mediante la información que el proceso tiene de su propia ejecución, sin importar la naturaleza de lo que realmente se está procesando.

La Figura 5 muestra los elementos principales relacionados con la conciencia de ejecución de un proceso.

```

Pensamiento={
  Proceso,
  Motivación
  Dedicación_Requerida,
  Dedicación_Asignada,
  Condiciones,
  Orden_Incremento_Dedicación,
  Fundamento={
    Conceptos,
    Recursos},
  Utilidad={
    Conceptos,
    Deseo={
      Modelo_Satisfacción,
      Satisfacción_Actual,
      Expectativa_Satisfacción}},
  Modelo_Dinámica,
  Modelo_Dedicación}

```

Figura 5: Conciencia del proceso

Un pensamiento se construye asociado a un concepto especial que tiene el rol de Deseo.

Cuando el deseo se visualiza, un pensamiento apropiado se construye y activa. El motivo del pensamiento es a partir de ese momento satisfacer dicho deseo.

Para ello el pensamiento tiene asociado un proceso.

La información de entrada para que el proceso puede avanzar en su ejecución está definida mediante un conjunto de conceptos que tienen el rol de Fundamento.

El pensamiento tiene conocimiento sobre la dinámica del problema que el proceso está resolviendo. Tiene también conocimiento sobre los requerimientos de seguridad y las necesidades de procesador para

satisfacer dichos requerimientos en cada una de las condiciones posibles. Y, a través del deseo, sobre las situaciones que lo satisfacen y con qué grado.

Todo este conocimiento se define mediante tres modelos explícitos MoDin, MoDed y MoSat, que permiten construir una interfaz de negociado de dedicación con el sistema de atención.

3.3.1 Emociones intrínsecas de los pensamientos y organización mental

Las emociones pueden ser definidas por la aplicación para que el agente responda ante situaciones específicas en su entorno.

Interesan sin embargo, en este trabajo, otro conjunto de emociones que llamamos intrínsecas, que son independientes de la aplicación y que se crean y asocian a cada uno de los pensamientos activos.

Estas emociones intrínsecas contemplan el hecho de pensar más que en lo que realmente se está pensando. Tienen que ver con las dependencias entre los pensamientos y los conceptos sobre los que estos trabajan. Así, un sistema puede alterar su ánimo en función de las expectativas de éxito, independientemente de cuál es la naturaleza del éxito o fracaso.

Los deseos se formulan con un nivel de motivación (importancia), y los pensamientos que los satisfacen heredan dicha motivación como su motivación básica. Esta motivación básica se gestiona con la emoción intrínseca asociada al deseo. Otras emociones intrínsecas asociadas a la confianza en el fundamento, a la urgencia y oportunidad en la satisfacción del deseo, modifican las motivaciones básicas para reorganizar la secuencia de ejecución.

Cada pensamiento aplicado es un Pensamiento Seguro, en el sentido de que contempla la dedicación mínima necesaria para satisfacer los requerimientos críticos de seguridad.

Pueden darse distintos esquemas de relación entre pensamientos y los conceptos asociados. En función de cómo ciertos conceptos participan como entradas o salidas en dichos pensamientos. Figura 6.

Pensamientos Independientes
P1(D:C1,F:C3)
P2(D:C2,F:C4)
Pensamientos Dependientes – Recursos Compartidos
P1(D:C1,F:C3)
P2(D:C2,F:C3)
Pensamientos Dependientes – Recursos Encadenados
P1(D:C1,F:C2)
P2(D:C2,F:C3)

Px = Pensamiento x Cx = Concepto x D: = Concepto de Salida de Pensamiento (Deseo) F: = Concepto de Entrada de Pensamiento (Fundamento)

Figura 6: Esquemas de pensamientos

EMOCIONES SOBRE EL DESEO

La emoción intrínseca básica de todo pensamiento es aquella que lo motiva en función de la importancia del deseo que dicho pensamiento está intentando satisfacer.

Los deseos se formulan con un nivel de importancia. El pensamiento que automáticamente es creado cuando el deseo se visualiza, hereda una motivación relacionada con la importancia del deseo.

El deseo puede modificar su importancia durante el ciclo de vida de su pensamiento asociado. Puede valorarse como más importante o como menos importante en un ciclo de atención dado. Entonces, la emoción intrínseca del sistema asociada al deseo modifica la motivación básica del pensamiento.

EMOCIONES SOBRE EL FUNDAMENTO

El pensamiento es un proceso. Los datos de entrada sobre los que el proceso opera disponen de valoraciones de confianza de aplicabilidad.

La emoción intrínseca asociada al fundamento modifica (modula) la motivación básica del pensamiento en función de la confianza del fundamento.

Por ejemplo, un pensamiento asociado a la ejecución de un plan puede desmotivarse mientras la confianza en el plan haya decaído en un momento dado y disminuyan entonces las expectativas de éxito. Puede también motivarse cuando dicha confianza aumente (seguramente porque otro pensamiento activo de diseño se haya completado).

EMOCIONES SOBRE LA DEDICACIÓN

Aun estando motivado por la satisfacción de un deseo importante y disponiendo de confianza en la información de partida, el pensamiento puede motivarse o desmotivarse en base a la expectativa de éxito.

Esta expectativa tendrá relación con la dedicación mental (dedicación de procesador).

La dedicación del procesador necesaria está relacionada con las condiciones del problema y los requerimientos estrictos de seguridad. Entonces, dada una dedicación ofertada por el sistema de atención, el proceso debe ajustar las condiciones para satisfacer los requerimientos de seguridad.

Dadas estas condiciones del problema (condiciones seguras) el modelo de dinámica junto con el modelo de satisfacción pueden dar una expectativa de satisfacción esperada.

Entonces, la emoción intrínseca asociada a la dedicación y expectativa de satisfacción puede provocar una segunda modulación de la motivación base del pensamiento.

ACTITUD: CARÁCTER Y HUMOR

No hay restricciones iniciales en cuanto al sentido e intensidad de las modulaciones de motivación que pueden realizar las emociones intrínsecas.

La definición concreta está asociada a la actitud del agente, definida en cada momento por el carácter y por el estado de humor.

3.4 ATENCIÓN

El *Sistema de Atención* organiza la ejecución de los procesos, asignando una dedicación de procesador a cada uno de ellos y realizando el cambio de contexto entre los mismos. Para ello realiza un negociado con los pensamientos, ya que éstos disponen de información relevante para garantizar su propia ejecución de forma segura o para determinar el grado de satisfacción de sus deseos (requerimientos funcionales).

La atención se produce al ejecutar un proceso intrínseco del sistema, que está en ejecución de forma permanente (alta prioridad). El Proceso periódico de Atención debe:

- Actualizar la confianza en los fundamentos
- Realizar las evaluaciones emocionales
- Actualizar las motivaciones de los pensamientos
- Planificar la dedicación mental
- Iniciar y suspender procesos

El coste del Proceso de Atención es función de la carga mental: número de conceptos, de pensamientos y de emociones, así como de su complejidad.

Cada uno de estos objetos conscientes es capaz de ofrecer una estimación de su coste de ejecución al sistema de atención, de forma que éste puede determinar su carga total.

El periodo de ejecución está relacionado con la dinámica del problema de aplicación. En una implementación sencilla, todas las actividades del proceso de atención se realizan en cada Ciclo de Atención. En una implementación más compleja, la atención puede distribuirse en el tiempo en función de las dinámicas particulares de cada uno de los subproblemas de aplicación. Esto sin embargo obliga a realizar previsiones más complejas sobre la carga total del sistema de atención, lo cual supone otra carga extra.

En cada ciclo de atención se cuestiona qué se está pensando, y se decide sobre lo que pensar.

Cuando el sistema de atención planifica la dedicación, negocia con cada pensamiento para adaptar sus expectativas a la capacidad mental disponible. En este negociado se cumplen estrictamente los requerimientos de seguridad y se optimizan los requerimientos funcionales, maximizando la satisfacción de los deseos que los pensamientos están intentando alcanzar. Para ello, cada pensamiento actualmente atendido tiene derecho de mantener su dedicación hasta conseguir llevar al sistema a unas condiciones seguras para la otra dedicación que se le está ofertando.

En ese sentido, el sistema de atención es quien decide sobre la dedicación a asignar, pero facilita recursos a los pensamientos para que adapten sus condiciones y puedan satisfacer siempre sus requerimientos de seguridad.

INTERFAZ DE NEGOCIADO DE ATENCIÓN

Para realizar el negociado de atención se necesita que los pensamientos y el proceso de atención implementen una interfaz:

Atención → Pensamiento	Pensamiento → Atención
Motivación?	Dedicación?
Dedicación(Satisfacción)?	Dedicación!?
Satisfacción(Dedicación)?	
Tiempo(Dedicación!)?	
Δ Dedicación(Tiempo)?	
Dedicación!	

(?) Pregunta (!) Orden

Figura 7: Interfaz de negociado de atención

ALGORITMO DE NEGOCIADO DE ATENCIÓN

Durante el negociado de atención se mantienen dos listas principales de pensamientos.

- Dedicación_Actual (lista de pensamientos actualmente atendidos)

- Dedicación_Planificada (lista de pensamientos que se está planificando atender)

Otras dos listas auxiliares muestran los pensamientos que perderán (que deben ceder) dedicación y los pensamientos que ganarán dedicación.

El negociado se realiza sobre estas listas, que ordenan los pensamientos de mayor a menor motivación. Las motivaciones de los pensamientos son generadas por el sistema emocional, y pueden ser resultado de evaluaciones emocionales con distinto grado de objetividad. Ciertas emociones basadas en evaluaciones que podríamos considerar subjetivas (urgencia, oportunidad, etc.) pueden dar como resultado cambios en la efectividad del sistema ante ciertos tipos de problemas.

Sobre la lista de planificación de dedicación, para optimizar la satisfacción de los requerimientos funcionales, el negociado se realiza en dos fases.

En la primera fase, se asigna dedicación de procesador a los pensamientos más motivados para alcanzar un nivel de satisfacción mínimo, definido en un Modelo de Atención (que relaciona la motivación con la satisfacción mínima esperable) Este modelo de atención depende del carácter del agente. Este proceso de reparto se detiene cuando se alcanza un umbral de dedicación en el procesador (este umbral es definido también por el carácter del agente y su estado de ánimo general).

Las listas auxiliares de pensamientos que pierden o ganan dedicación se actualizan en este proceso.

Cuando se ha determinado el número de pensamientos que podrían ser mantenidos por el procesador, se determina si hay pensamientos en la lista de pérdida de dedicación. Si no hay ninguno, se realiza una segunda fase de asignación de procesador en la que se intenta alcanzar la satisfacción máxima posible desde los pensamientos más motivados a los menos. De nuevo hasta agotar la capacidad del procesador. Y la lista de planificación se le pasa al despachador de procesos para que la convierta en la lista de dedicación actual.

Si por el contrario sí hay pensamientos que pierden dedicación, la lista de atención que se está planificando se recorre de menor a mayor motivación, buscando pensamientos que aparezcan en la lista de aumento de dedicación, esta dedicación se recorta en la cantidad necesaria para mantener los pensamientos actualmente atendidos.

La lista de pensamientos que pierden dedicación es utilizada para transmitir peticiones de ajuste de

dedicación a cada uno de los pensamientos involucrados. Éstos son avisados para que adapten sus condiciones a la dedicación que se les está ofertando. Cada pensamiento, mediante la aplicación de sus modelos de dinámica y de dedicación es capaz de informar sobre el tiempo requerido para éste ajuste.

3.5 RELACIÓN

Finalmente, el Sistema de Relación comunica al agente con su entorno.

3.5.1 Planos Conscientes

Los conceptos que se manipulan en los pensamientos pueden ser entradas o salidas a éstos. Ciertos conceptos se manipulan en más de un pensamiento, entonces son considerados Recursos para gestionar la utilización simultánea.

Un mismo tipo de concepto puede definirse en distintos planos conscientes, que representan mundos virtuales utilizados para la previsión en el diseño de planes.

Un plano especial, y único, denominado Plano Real, está asociado a los conceptos relativos al sistema de relación. Los conceptos asociados a los dispositivos efectores son bloqueados y liberados por los pensamientos que los utilizan, y los conceptos asociados a los dispositivos sentidos pueden ser compartidos por distintos pensamientos.

4 CARÁCTER DEL AGENTE

Dada una situación en el entorno, la respuesta del agente RTEA depende sensiblemente de su sistema afectivo.

Por una parte, las emociones dependientes de la aplicación consideran las situaciones frente a las que el agente no debería ser indiferente. Dependiendo del tipo de problema y de los objetivos que se quiere que el agente alcance, se contemplarán unas situaciones u otras, y se ajustará la valoración y la respuesta emocional del agente ante las mismas.

Así, en una aplicación de robótica móvil, por ejemplo, las posiciones del robot y de otros sólidos próximos, y sus velocidades, puede ser evaluada emocionalmente y provocar un estado de miedo a colisión. Este estado emocional, motiva claramente respuestas de evitación de colisión.

Por otra parte, en el caso de las emociones del sistema, lo que se pretende es considerar la situación mental del agente, y que ésta tenga también

influencia en la respuesta. Situaciones como la excesiva prolongación de un pensamiento sin alcanzar la satisfacción del deseo, la percepción de la situación de oportunidad de poder finalizar de forma inminente un pensamiento y liberar la mente de carga excesiva, etc., pueden ser evaluadas y provocar sensibles reorganizaciones de la atención mental.

Al ser estas situaciones mentales tratadas de forma general, independientemente de la aplicación, las emociones intrínsecas se construyen de forma automática para cada uno de los pensamientos que se activan.

4.1 QUÉ ES EL CARÁCTER Y DÓNDE SE DEFINE

No es obvio el efecto final que puede provocar un conjunto de emociones y su ajuste sobre el nivel de satisfacción de los requerimientos funcionales.

Es por tanto importante la definición de distintos caracteres de agente, constituidos por conjuntos de emociones intrínsecas y los ajustes de las mismas.

Los distintos caracteres de agente pueden ponerse a prueba sobre distintos tipos de problema y determinar qué caracteres se adaptan bien o mal a qué problemas.

El carácter queda entonces definido en forma de valoraciones y efectos emocionales que se construyen de forma automática para cada pensamiento que se activa.

Adicionalmente el carácter se define también en el sistema de atención, donde los criterios de asignación de dedicación establecidos por el Modelo de Atención, pueden también ajustarse y catalogarse como distintos tipos de caracteres.

4.2 CÓMO AFECTA EL CARÁCTER EN LA ORGANIZACIÓN MENTAL

La influencia del carácter en la organización mental se evidencia en la modificación de las motivaciones de los pensamientos que producen las emociones. Ya que la sensibilidad emocional y la respuesta emocional se ajusta en el carácter.

Dada una ordenación de los pensamientos en base a su motivación, la asignación de dedicación sigue también un modelo elegido según un carácter.

Así, definiendo distintos niveles de contribución y efecto emocional, y definiendo distintos modelos de asignación de dedicación, disponemos de distintos caracteres de agente, que pueden dar distintas

prestaciones funcionales en función del tipo de problema.

4.3 EJEMPLO

Consideremos un estado emocional de excitación asociada a una valoración sobre la urgencia de un pensamiento.

VALORACIÓN EMOCIONAL DE URGENCIA

La motivación de un pensamiento debería estar directamente relacionada con la importancia del deseo que dicho pensamiento está intentando satisfacer.

¿Por qué los agentes reales se dejan influir por la sensación de urgencia y dedican su tiempo y sus recursos a realizar actividades consideradas, en principio, menos importantes que otras que quedan temporalmente suspendidas pese a su importancia?

De alguna manera está manifestándose la sensación de incertidumbre que el agente tiene sobre lo que acontecerá en el futuro.

Hay una gran limitación en nuestros modelos mentales sobre el mundo a la hora de determinar qué recursos serán realmente necesarios en el futuro. Cuando el plazo es mayor, la incertidumbre es mayor también, y parece que es poco práctico afinar planes a largo plazo, siendo suficiente con considerar planificaciones aproximadas que se afinan en el entorno temporal más cercano, donde los modelos de previsión pueden ofrecer mayores garantías.

La urgencia es una valoración emocional que contempla las capacidades (recursos), niveles de satisfacción de objetivos y plazos. La valoración contempla estos conceptos y detecta combinaciones de los mismos en la que los plazos son insuficientes para los recursos disponibles y la satisfacción deseada. Todo esto basándose en los modelos de dinámica, de dedicación y de satisfacción disponibles.

El estado emocional derivado depende del carácter del agente, que define precisamente la valoración.

Alcanzado un nivel de excitación debida a la urgencia, el sentido y la intensidad de la respuesta también depende del carácter.

Un carácter tranquilo percibiría la urgencia con el plazo más agotado que con el que lo haría un carácter nervioso.

Un carácter con una escala de valores estricta sería imparcial con el conjunto de pensamientos y se

dejaría influir poco, con lo que los pensamientos obtendrían dedicación según la importancia de los deseos subyacentes. Un carácter con una escala de valores menos estricta se dejaría influir.

Tipo Carácter	Valoración Emocional	Estado Emocional	Tipo Carácter	Tipo Carácter	Motivación Pensamiento Urgente
Nervioso	Muy Sensible	Excitación ante Urgencia	Muy Influyente	Optimista	Positiva
				Pesimista	Negativa
Tranquilo	Poco Sensible		Poco Influyente		Indiferente

Figura 8: Ejemplo de influencia del carácter ante la valoración emocional de urgencia

Dependiendo de si el carácter es considerado positivo o negativo los efectos serían sensiblemente distintos. Un carácter positivo motivaría al alza el pensamiento urgente, como dándole una importancia (aparente) mayor al deseo urgente.

Esta concepción positiva de la vida ve en el largo plazo del resto de pensamientos menos urgentes (aunque posiblemente más importantes), un espacio donde puede haber de todo, donde se tendrán oportunidades para avanzar los pensamientos importantes cuando sea necesario, donde se podrá hasta replantear la escala de valores, etc. Entonces, es mejor 'liquidar' este pensamiento, que por ser urgente, pasará pronto y liberará recursos, y obtener un rédito de satisfacción positivo inmediato, y confiar en que el futuro proveerá los recursos necesarios.

Un carácter negativo, al contrario, motivaría a la baja el pensamiento urgente, precisamente por una concepción pesimista sobre el futuro.

Muchos impedimentos pueden aparecer en el futuro incierto que dificultarían la consecución de los objetivos realmente importantes. Si este pensamiento urgente, lo es porque no hay plazo, y no va a satisfacer el nivel deseado, ¿sería quizás mejor renunciar y no comprometer los objetivos realmente importantes.

5 CONCLUSIONES

Se han presentado las características básicas de la arquitectura de agente RTEA.

Una parte importante de la arquitectura es la regulación emocional, y en concreto la regulación producida por las denominadas emociones intrínsecas de los pensamientos, que son independientes del problema.

En esta fase del proyecto se están definiendo conjuntos de emociones intrínsecas que definan distintos caracteres de agentes, que son de interés en distintos tipos de aplicación.

Agradecimientos

Este trabajo ha sido parcialmente financiado con cargo al proyecto DPI2002-04432-C03-01.

Referencias

- [1] Arkin, R.C., Ali, K., Weitzenfeld, A., and Cervantes-Perez, F., (2000), "Behavioral Models of the Praying Mantis as a Basis for Robotic Behavior", *Journal of Robotics and Autonomous System*, 32 (1):39-60.
- [2] Arkin Ronald C., Fujita M., Takagi T. and Hasegawa R. (2003) "An Ethological and Emotional Basis for Human-Robot Interaction", *Journal of Robotics and Autonomous Systems*, 42 (2003) 191-201.
- [3] Arkin, R.. (2004) "Moving Up the Food Chain: Motivation and Emotion in Behavior-Based Robots", In *Who Needs Emotions? The Brain Meets the Robot*. Editors: Jean-Marc Fellous and Michael A. Arbib To be published by Oxford University Press in Dec. 2004.
- [4] Breazeal C., (2003), "Emotion and Sociable Humanoid Robots". E. Hudlika (ed), *International Journal of Human Computer Studies*, 59, pp. 119-155.
- [5] Domínguez C., Hassan H., Crespo A, (2004). "Real-Time Robotic Agent Based On Emotions". *International Symposium on Robotics & Applications*. World Automation Congress 2004.
- [6] Euron (2004). *European Robotics Research Network*. "Euron Research Roadmaps", Key Area 1 on Research Coordination, 26 March 2004. <http://www.euron.org>
- [7] Moshkina, L., and Arkin, R.C., (2003), "On TAMEing Robots", *Proc. IEEE International Conference on Systems, Man and Cybernetics*, Oct. 2003.
- [8] Watson, D., 2000, "Mood and Temperament", New York, The Guilford Press.