

Aprendiendo a Aprender: de Máquinas Listas a Máquinas Inteligentes

Bogdan Raducanu¹ y Jordi Vitrià^{1,2}

¹ Centre de Visió per Computador, Edifici O - Campus UAB

² Departament de Ciències de la Computació, Universitat Autònoma de Barcelona
08193 Bellaterra, Barcelona

España

E-mail: {bogdan, jordi}@cvc.uab.es

Abstract. Desde su aparición, hace más de cinco décadas, uno de los retos más ambiciosos de la Inteligencia Artificial ha sido la creación de sistemas computacionales inteligentes. A pesar de los impresionantes avances alcanzados, todavía estamos lejos de tener máquinas que se acerquen al nivel de la inteligencia humana. Esto se debe al hecho de que los sistemas de hoy en día carecen de sentido común, un elemento característico de las personas. El presente artículo está enfocado sobre el análisis de las distintas estrategias de aprendizaje, tipos de contexto que intervienen en el proceso de aprendizaje y su aplicación en la interacción persona-máquina.

Keywords: sistemas inteligentes, aprendizaje cognitivo, contexto, robótica social, reconocimiento de caras

1 Introducción

Conocimiento no es equivalente a inteligencia. Hoy en día (debido a los avances en Inteligencia Artificial y potencia de cálculo), las máquinas son capaces de hacer cosas remarcables: hay algoritmos de jugar ajedrez a nivel de grandes maestros, aplicaciones complejas para coordinar el desplazamientos de los efectivos militares en los campos de batalla, herramientas de diseño que nos ayudan desde el desarrollo de los circuitos electrónicos más sofisticados hasta las aeronaves más complejas. Nadie pone en duda todas estas evidencias. Pero en cambio, a pesar de la complejidad de los sistemas mencionados, ninguno de ellos es capaz de, por ejemplo, interpretar una fotografía, comentar un texto, contestar a una pregunta, cosas que son obvias para cualquier persona. A simple vista, no podemos negar que no existen las herramientas necesarias o que no hayan programadores capaces de desarrollar sistemas que presenten un cierto nivel de razonamiento (hay un número sin fin de ejemplos en el ámbito de los sistemas expertos). De hecho, el campo de Inteligencia Artificial está lleno de teorías y algoritmos de aprendizaje. La pregunta que nos planteamos entonces es: ¿qué componente, a parte del conocimiento especializado, se ha omitido en la fase de programación, para que una máquina sea realmente inteligente? La respuesta a la pregunta

anterior viene dada por el concepto de 'sentido común'. Pero que es el sentido común? La siguiente definición viene dada por Marvin Minsky en [11]: "the mental skills that most people share". Para nosotros, el sentido común es una característica tan natural, que muchas veces, en la vida cotidiana, lo ignoramos. En cambio, una máquina no tiene ni idea de como somos, que sentimos o cuales son nuestras preferencias.

Un sistema experto, por ejemplo, tiene una representación del conocimiento específico (premisas y reglas de inferencia) para solucionar un problema particular. Pero para dotarlo de sentido común, hace falta alimentarle con información mucho más general: conocimientos de física para saber como se comportan los objetos, conocimiento social para entender como interaccionan las personas, conocimientos de psicología, para entender como funciona la mente humana, etc. Para cada categoría de conocimiento, las personas utilizamos distintos modos de representación del mismo y distintas estrategias de razonamiento. Por lo tanto, dotar una máquina de sentido común no es especificarle que razonamiento aplicar sobre un cierto conjunto de datos (esto ya está solucionado), sino como seleccionar la estrategia de razonamiento adecuada y como seleccionar sobre que subconjunto de conocimiento (de la multitud de datos representados) debe aplicarla. Las complicaciones que aparecen aquí son aún mas grandes. En primer lugar, dado un conjunto tan grande de conocimiento general (hechos), ¿cuál es el modelo de representación adecuado? ¿Es posible que el cerebro humano utilice varias representaciones para el mismo hecho? En segundo lugar, tenemos que disponer de métodos para definir y representar el conocimiento funcional: las distintas estrategias de razonamiento, planificación, predicción, etc. Una discusión más detallada sobre estos aspectos está fuera del proposito de este trabajo, pero se puede encontrar en [17].

El artículo está estructurado de la siguiente manera: en la sección 2, hacemos un repaso de las teorías de aprendizaje existentes. La sección 3 está dedicada a la presentación del papel del contexto en el proceso de aprendizaje. En la sección 4 hablaremos de un estudio que está en marcha, de la aplicación del aprendizaje cognitivo en el entorno de la robótica social (como exponente del área de interacción persona-máquina). Finalmente, sección 5 contiene nuestras conclusiones y direcciones de trabajo futuro.

2 Teorías de aprendizaje

El término de aprendizaje, dentro de la Inteligencia Artificial, se refiere a la habilidad de una máquina de adquirir ciertos conocimientos de tal manera que, debido a los cambios en la representación interna de los datos (como consecuencia de nuevas experiencias), mejora su funcionamiento/comportamiento con el tiempo. Las preguntas que nos podríamos plantear son las siguientes: ¿Porque queremos que las máquinas aprendan? ¿Porque no diseñamos desde un principio, una máquina capaz de tener la funcionalidad deseada? Las respuestas a estas preguntas vienen desde varias direcciones: algunos conocimientos solo se pueden adquirir basandonos en ejemplos, puede haber una relación estricta en-

tre los datos de entrada y salida (o que los datos presenten una correlación oculta), el entorno de trabajo de la máquina cambia con el tiempo o incluso nuevos conocimientos son adquiridos por parte de los programadores. El aprendizaje converge de varias áreas: estadística, modelos cognitivos, teoría de control adaptativo, modelos psicologicos, etc. Más detalles sobre estos temas se pueden encontrar en [13].

En conclusión, y tal como vamos a ver con más detalles a continuación, hay una relación muy estrecha entre la adquisición de los datos, representación del conocimiento y estrategia de aprendizaje.

2.1 Aprendizaje basado en representación simbólica

En la metodología tradicional de la Inteligencia Artificial, la resolución de problemas se basaba en la decomposición funcional [8] y abstracción de los datos mediante una representación simbólica [12]. Los programadores partían de un conjunto de premisas e intentaban construir un modelo del mundo. Cuando se solicitaba una respuesta por parte del sistema, se intentaba encontrar una correspondencia uno-a-uno entre los datos de entrada y el motor de inferencia, que constituía la base de conocimientos del sistema. Por lo tanto el razonamiento se basaba en el ciclo 'percibir-interpretar-responder'.

El error que se cometía con la abstractización de los datos era que se utilizaban unos símbolos que no tenían nada que ver con el mundo real. Estos símbolos existían únicamente en la visión del programador. Con la abstracción de los datos, para el programador era imposible prever todos los casos posibles que podrían aparecer durante el funcionamiento del sistema. La explicación era que de este modo, se podría reducir sustancialmente los problemas a resolver. La arquitectura para este tipo de sistemas, era centralizada (de aquí la sintagma "brain-in-a box") y por lo tanto ocurría muy a menudo que el sistema quedaba bloqueado, cuando recibía un estímulo cuya respuesta no estaba prevista en su diseño. Un ejemplo clásico es representado por el MYCIN [16], un sistema experto para el diagnóstico de las infecciones bacterianas. El sistema no tenía ningún modelo sobre lo que es una persona o que le puede pasar. Si le decías por ejemplo que el paciente se ha hecho un corte y pierde sangre, el sistema intentaba determinar la causa bacteriana que pudo causar este problema.

La explicación para el fracaso de los sistemas basados en representación simbólica podría ser que el cerebro humano representa la información no solamente por su categoría, sino también por la modalidad con la cual ha sido obtenida. En otras palabras, la representación está basada también en los mecanismos sensoriales que han contribuido a la adquisición de la información [5].

2.2 Aprendizaje basado en comportamientos

Visto el aparente fracaso en el desarrollo de sistemas inteligentes mediante la abstracción del conocimiento, era obvio que un nuevo paradigma era necesario. En [4], el autor propuso el desarrollo de sistemas inteligentes que no necesitaban una representación centralizada del mundo. La idea era de que el aprendizaje

pueda tener lugar por el emparejamiento directo entre los datos de entrada (estímulos) y sus respuestas (el llamado "sensation-to-action"). En este caso, la arquitectura general del sistema era distribuida, representada por varios módulos que desarrollaban un comportamiento muy sencillo. Los comportamientos de alto nivel se realizaban como consecuencia de la combinación de estos comportamientos simples. Esta idea también ha sido utilizada por Minsky en [11] para explicar la emergencia de la inteligencia humana. En su visión, la mente está formada por un conjunto de 'agentes' que compiten y cooperan entre ellos.

Al ser excluida la representación explícita de los datos, no se puede hablar de un método para medir directamente la capacidad de conocimiento del sistema. En cambio, el nivel de aprendizaje se evalúa por el análisis de su comportamiento. En otras palabras, se puede decir que el sistema aprende cuando muestra un cambio en su comportamiento. Como consecuencia de ello, un requisito fundamental para estos sistemas, es representado por su instalación directa en el entorno donde iban a funcionar. Esto es debido al hecho de que su comportamiento se desarrolla basándose en la interacción directa con el entorno. Detrás de ello se encuentra la idea filosófica sobre el dualismo cuerpo-mente. Este es un concepto muy conocido en el campo de la Inteligencia Artificial y se refiere al hecho de que la mente no puede imaginar más allá de lo que está percibido por el sistema sensorial. Esta realidad está metafóricamente descrita en [10] por el sintagma: "the body is the anchor of the mind".

2.3 Aprendizaje cognitivo

El modelo de aprendizaje presentado en la sección anterior tiene una desventaja: está limitado en gran parte solo a los cambios del entorno, y en menor medida a la interacción entre los componentes internos. Por lo tanto, a veces es difícil de establecer si un cambio interno se ha debido realmente a la modificación de las condiciones externas. Por otro lado, hay sistemas cuyas características se parecen muchísimo a los descritos anteriormente, pero de los cuales no se puede decir que son inteligentes. Como contraejemplo, se puede mencionar un sistema cibernético. Consideremos un sistema de aire acondicionado. Es verdad que debido a la arquitectura interna (el termostato) este sistema cambia su comportamiento como consecuencia de la modificación de las condiciones externas (temperatura). Pero está muy lejos de ser considerado un sistema inteligente.

Por lo tanto, los investigadores han llegado a la conclusión de que el aprendizaje no se puede relacionar únicamente con los cambios comportamentales, sino que es también necesario estudiar los cambios internos que tienen lugar en la representación del conocimiento [14]. Con este nuevo paradigma, el aprendizaje puede tener lugar sin que se note un cambio aparente en el sistema. El aprendizaje cognitivo consiste en un proceso integrado, recursivo, que tiene como fin construir un modelo del mundo y una adaptación continua de este modelo. Por lo tanto, es el mismo sistema el responsable de como analizar, interpretar y modelar el mundo. El sistema aprenderá nuevos conceptos (desarrollará nuevas competencias) basándose en el conocimiento aprendido y en la experiencia adquirida. Con la llegada de nueva información, esta deberá ser analizada

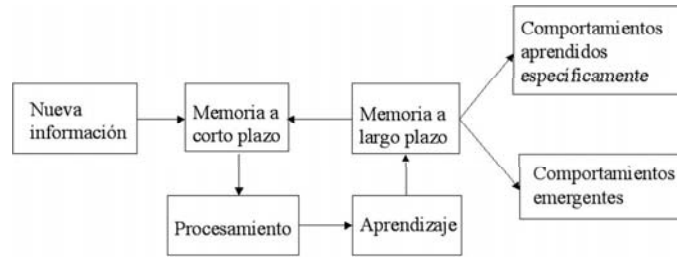


Fig. 1. Modelo de aprendizaje cognitivo. Para más detalles, ver texto.

y el sistema tomará la decisión de, si es información útil, integrarla en la representación existente. A veces, se puede incluso llegar a cambiar la estructura de la representación como consecuencia de esta nueva información. Como respuesta, el sistema podrá desarrollar dos clases de comportamiento: una clase de comportamientos aprendidos específicamente y otra clase de comportamientos emergentes. Este modelo de aprendizaje cognitivo está representado en la figura 1.

Una característica muy importante cuando hablamos del aprendizaje cognitivo es el acceso a la información acumulada. Hay que destacar dos tipos de memorias: una a corto plazo y otra a largo plazo. La memoria a corto plazo se refiere a la información que se necesita para finalizar una tarea específica (con un propósito y una duración muy bien definidas). Después, la información se puede eliminar (sin que afecte de algún modo la viabilidad) para no ocupar innecesariamente los recursos del sistema. Por otro lado, la memoria a largo plazo es la información necesaria que le permite al sistema existir y funcionar en el tiempo (medido en años por ejemplo). Un problema que puede ocurrir con la memoria a largo plazo, es el proceso de 'olvido' o de 'degradación' (perdida parcial de las características asociadas con una cierta información). El fenómeno de olvido es una propiedad fundamental para el funcionamiento del cerebro humano.

Un elemento muy esencial en el proceso de aprendizaje cognitivo está representado por el contexto: el conjunto de factores que son determinantes en la adquisición, representación y (más tarde) recuperación de la información. En el siguiente párrafo vamos a introducir la noción de contexto y como se representa en distintos áreas.

3 Contexto

Según [23], el término de 'contexto' tiene sus fundamentos en lingüística. La palabra viene compuesta por las partículas 'con' y 'texto' y se refiere al significado extraído de la lectura de un texto. Hoy en día, la acepción del término es mucho más amplia, y se refiere a un marco particular en el proceso de comunicación, basado en unas experiencias comunes.

Con el paso del tiempo, la utilización del término 'contexto' se ha especializado en distintas áreas de conocimiento. En [3] se puede encontrar una discusión

muy detallada sobre este tema. A continuación, vamos a destacar la utilización del 'contexto' solo en las áreas de interés.

En Inteligencia Artificial, viene asociado con el 'frame problem': representar en el lenguaje de la lógica el resultado de una acción de forma implícita, sin recurrir a una representación explícita de los efectos que no se han producido como consecuencia de la acción. Es uno de los problemas más difíciles con el cual se ha confrontado la Inteligencia Artificial.

En el ámbito de la comunicación (escrita o verbal), el 'contexto' se refiere a unas propiedades del proceso de interacción entre varios agentes. Esta interpretación es opuesta a la del 'contexto' como propiedades de un fenómeno en particular. En otras palabras, la noción de 'contexto' no existe sin interacción. El 'contexto' es considerado como una 'historia' de todo lo que pasó en un cierto periodo de tiempo (desde que se inició la interacción), el conjunto de conocimientos de que disponen los agentes y de las particularidades derivadas del tema sobre el cual se están enfocando en un cierto instante. De este modo, el contexto aparece como un 'espacio de conocimiento compartido'. Como consecuencia, se pueden desarrollar herramientas (asistentes) cuyo fin es predecir y anticipar las solicitudes de los agentes. En el caso de la comunicación escrita, los ejemplos vienen dados por los asistentes asociados a los editores de texto: nada más introducir las primeras letras de una palabra, el asistente busca en su diccionario las posibles opciones para poder continuar. En el caso de la comunicación verbal, teniendo información sobre el tema de la conversación podemos eliminar la incertidumbre provocada por una expresión ambigua, implícita de los hechos. Por ejemplo, en la siguiente frase: 'He visto un gato por la calle', al no conocer el contexto, no se puede inferir el significado. No se sabe si la persona se refiere al gato como animal, o al gato como herramienta para coches.

En el ámbito de la visión, el término 'contexto' se refiere tanto al entorno que rodea un objeto en particular, como a las propiedades intrínsecas del mismo objeto. Incluso hoy en día, a pesar de los avances realizados en el área de procesamiento de imágenes y el reconocimiento de patrones, la identificación visual de objetos queda un problema parcialmente solucionado. Visto desde el punto de vista de la Inteligencia Artificial, la tarea de describir una escena (en términos de los objetos que la componen) es uno de los hitos más ambiciosos que han quedado por resolver. Entre los factores que dificultan el reconocimiento visual se pueden mencionar las condiciones de iluminación, los cambios en la apariencia (vistas desde diferentes ángulos, deformaciones, transformaciones lineales), posibles oclusiones, etc. Por esta razón, el uso del contexto pretende simplificar el proceso de reconocimiento por parte del sistema visual. Con el uso del contexto, no solamente se puede conseguir la eliminación de muchos errores, sino también la ambigüedad que puede dificultar a veces la toma de decisiones.

Experimentos en análisis de escenas [2] han confirmado que el sistema visual humano utiliza de manera extensiva el contexto para facilitar la detección y el reconocimiento de objetos. Además, en el mundo real existe una relación muy estrecha entre un objeto y el entorno en el cual está situado. Por lo tanto, la decisión sobre la presencia o la ausencia de un objeto en la escena está en

gran medida influenciada por ella; la presencia de distintos tipos de objetos puede estar fuertemente correlacionada: por ejemplo, si se consigue detectar un monitor en una imagen, se puede esperar a que se encuentre también un teclado. Parece que el sistema visual lo primero que hace es un análisis global de la escena para poder estimar los objetos que pueden aparecer en ella. El contexto puede ayudar a la identificación de los objetos presentes en la escena de dos modos [18],[19]: en primer lugar, cuando las características de los objetos son parcialmente observables o están afectadas por ruido; en segundo lugar, asumiendo que la identificación se ha producido satisfactoriamente, el contexto todavía puede ayudar a la hora de eliminar las posibles incertidumbres creadas en la fase de clasificación del objeto. Además, los mismos autores proponen una representación del contexto teniendo en cuenta las características generales que aparecen en la imagen (realizar una representación de la imagen con un número reducido de dimensiones, por ejemplo). Esta transformación se puede realizar de una manera relativamente sencilla, sin la necesidad de identificar unas regiones en concreto en la imagen. Con esta representación luego, se facilita la detección de objetos individuales, porque el contexto puede dar indicios muy valiosos sobre la posición y el tamaño de los objetos en la imagen.

4 AiboFace: un estudio para el desarrollo del aprendizaje cognitivo en robots

Desde su creación, los sistemas computacionales han sido siempre enfocados hacia la máquina y no hacia las personas. Hasta hoy en día, la interacción persona-máquina suponía que el usuario debe comprender el funcionamiento de las máquinas, como ‘piensan’ y como están construidas. Como consecuencia, estamos obligados a trabajar en sus términos, utilizando su lenguaje y unos dispositivos específicos para comunicarnos con ellas (ratón, teclado, etc.). Con la aparición de la realidad virtual, la cosa era aún más absurda: estamos obligados a sumergirnos en un mundo sintético, creado por ellas.

Pero en el futuro, los investigadores auguran unos cambios fundamentales en la interacción persona-máquina. En el futuro, la interacción será centrada en el usuario. En otras palabras, el usuario no tendrá porque estar preocupado por la presencia de las máquinas o como funcionan. En cambio, las máquinas estarán previstas con capacidad sensorial, que les permita identificar la presencia de una persona y estar atenta en cada momento a sus acciones, pero al mismo tiempo respetando nuestra privacidad e intimidad. Como consecuencia, ellas estarán omnipresentes en nuestras vidas, rodeándonos en el trabajo y en el hogar. De hecho, esta realidad ha sido anticipada por Mark Weiser en [22].

Dentro de esta visión, un lugar destacado está representado por la robótica social. En la robótica social, los investigadores consideran los robots como unas plataformas para formalizar y probar las capacidades cognitivas de los humanos. Al mismo tiempo, por la implementación de estos modelos, podemos tener una mejor comprensión de su funcionamiento en las personas. Ultimamente, se intenta extender el uso de estos robots para convertirlos en verdaderos asistentes



Fig. 2. El robot AIBO de SONY con sus dos juguetes favoritos: la pelota y el hueso.

para personas de la tercera edad. Experimentos psicológicos y sociales han demostrado que pueden ser de gran ayuda en levantar el estado de ánimo de las personas [21], pero también de avisar a los servicios competentes en caso de una situación de emergencia. Por otro lado, se ha estudiado el uso de estos robots en clínicas pediátricas y las observaciones resultadas de la interacción con los niños han sido analizadas [24]. Debido a la gran complejidad que supone el desarrollo de esta clase de aplicaciones, varias áreas están implicadas: diseño de interfaces, psicología, neurociencias, etc.

Este nuevo paradigma en robótica, contrasta fuertemente con la visión que se tenía sobre los robots hace un par de décadas. Entonces, eran visto solamente como unos sustitutos para las tareas que implicaban acciones repetitivas o que suponían un alto nivel de riesgo para el operador humano. Los primeros robots que se fabricaron eran muy limitados, siendo diseñados para unos entornos muy específicos y siendo capaces de realizar unas tareas muy concretas.

En el caso de los robots sociales, una de las tareas básicas que tienen que solucionar es la detección de personas. Las caras representan de lejos el mejor indicio sobre la presencia de una persona en la vecindad del robot. El argumento para esta afirmación tiene sus raíces en la visión biológica. En [7], los autores reclaman el hecho de que los recién nacidos llegan al mundo con una pre-disposición de reconocer caras. Parece ser que, en general están atraídos por patrones en movimiento que tienen estructura parecidas a la cara. Otro argumento presentado en el mismo trabajo, subraya el hecho de que para los humanos nos es más fácil reconocer caras (si están presentadas en posición frontal) que cualquier otro objeto.

Para nuestro estudio, utilizamos un robot AIBO [1] de la marca SONY (figura 2).

Creado inicialmente con el propósito de servir como un juguete de entretenimiento más, ha sido rápidamente adoptado por la comunidad científica que ha visto en él uno de los mejores entornos para el desarrollo y prueba de las teorías del ámbito de la robótica social (como el aprendizaje cognitivo, por ejemplo). El nombre del robot puede ser interpretado de dos maneras: por un lado, la palabra 'aibo' significa 'compañero' en japonés; por otro lado, su nombre puede ser visto como la combinación entre Inteligencia Artificial (AI - por sus siglas en inglés) y 'roBOt'. AIBO es un robot que viene

pre-programado con capacidad para aprender, responder a una variedad de estímulos (reconocimiento de voz, tacto, visión), expresar deseos y mostrar emociones. Sin duda, AIBO es 'algo' distinto. Esto se nota desde el primer instante que se quiere adquirir uno. El personal de venta de SONY te advierte de que un AIBO no se compra, sino se adopta.

Debido a su curiosidad, el AIBO está en un proceso continuo de aprendizaje y adaptación al entorno [9], [15]. Como consecuencia de ello (a las particularidades con las cuales se han confrontado cada uno de ellos), podemos afirmar que no existen dos AIBOs iguales en el mundo (del mismo modo que no existen dos personas iguales).

A parte de los comportamientos pre-programados, el AIBO viene también con su propio entorno de desarrollo de aplicaciones (unas librerías de funciones para el Visual C++), llamado AIBO Remote Framework. Estas librerías nos permiten desarrollar aplicaciones del tipo cliente-servidor, entre el PC y el AIBO, teniendo como 'soporte' una conexión de red inalámbrica entre los dos. El problema que nos planteamos estudiar es el aprendizaje gradual por parte del AIBO de las personas, mediante el reconocimiento facial. Para este propósito necesitamos desarrollar unos algoritmos específicos. La aplicación que nos planteamos es relativamente nueva en el ámbito de reconocimiento de caras y se podría expresar como 'reconocimiento no-supervisado'. La idea es la siguiente: al principio, el robot no conoce ninguna persona (el conjunto de aprendizaje es vacío). Con el transcurso del tiempo, a medida que ve nuevas caras, empieza a construir de forma incremental la base de datos.

La estrategia de aprendizaje que se quiere emplear se puede dividir en dos fases, cada una de ellas automatizada. En una primera fase, al robot se le puede 'enseñar' caras, pero no es capaz de diferenciarlas. En una segunda fase, se pretende que el mismo robot desarrolle competencias para ser capaz de clasificarlas (del conjunto que el mismo ha adquirido). Durante este proceso de detección/reconocimiento, algunas características de las caras se podrán distinguir claramente de otras, mientras que algunas se inferirán del contexto. Esta estrategia de aprendizaje es muy similar a la utilizada por humanos. En los primeros meses de nuestra infancia, solo somos capaces de distinguir la clase 'cara' dentro del conjunto general de objetos [6]. Al cabo de un cierto periodo de tiempo, y a medida que nuestras capacidades cognitivas han evolucionado, no solamente somos capaces de identificar las personas, sino también reconocer el género (hombre/mujer) o diferenciarlas según la edad (niño/joven/adulto/anciano).

Al nivel más técnico, el sistema de reconocimiento empleará dos tipos de memoria: una memoria a corto plazo, y una memoria a largo plazo [25]. Con la memoria a corto plazo pretendemos que el robot sea capaz de mantener una coherencia sobre la identidad de la persona a quién esta viendo mientras dure la sesión (una sesión se define como el periodo transcurrido desde la aparición de una persona en la escena, hasta su salida). Con la memoria a largo plazo, el propósito es de crear una base de datos con los 'amigos' del robot (reidentificación: ¿de que me suena esta cara?). Entre los dos tipos de memoria existe una relación muy estrecha: las nuevas imágenes de cara tomadas durante las sesiones, se añadirán a la memoria a largo plazo, para actualizar de este modo su contenido. Este proceso de 'renovación' tendrá una componente adaptativo en el sentido de que cuando se construya el modelo de cara de una persona, las instancias más recientes tendrán un peso mayor en el cálculo del modelo, que las más viejas. Se llegará a un cierto momento, incluso, que las imágenes más antiguas sean completamente y definitivamente eliminadas de la base de datos (es la fase de 'olvido' del proceso de aprendizaje).

Además, cada persona tendrá asociada un 'peso' proporcional con el número de veces que ha sido vista. Con ello se pretende que el robot desarrolle un comportamiento

particular (que muestre 'más interés') hacia las personas que han sido vistas más a menudo (tienen un 'peso' grande), con respeto a personas que han sido vistas menos veces (tienen un 'peso' pequeño).

En estos momentos nos encontramos en la primera fase del proceso de aprendizaje descrito anteriormente: adquisición automática de caras. En concreto, hemos implementado un detector de caras basado en [20]. El detector está relacionado con los circuitos motrices de la cabeza del robot, de tal modo que el efecto conseguido es de un seguimiento activo de la persona: el robot mueve su cabeza en concordancia con el movimiento de la persona en su campo visual, de tal modo que la persona se queda siempre centrada en la imagen capturada por la cámara del robot. En figura 3 mostramos unas instancias del proceso de captura de caras.

5 Conclusiones y trabajo futuro

En este artículo hemos presentado las causas por las cuales los sistemas computacionales de hoy son solamente listos, sin ser inteligentes. El componente que falta en su diseño es el 'sentido común', un rasgo característico de las personas. Después de repasar brevemente las teorías de aprendizaje existentes, hemos subrayado la relevancia del contexto en el proceso de aprendizaje. Su uso es imprescindible si deseamos obtener buenas respuestas por parte del sistema incluso en situaciones cuando los datos son incompletos o corruptos. También su papel puede ser fundamental en el proceso de clasificación para la eliminación de la incertidumbre. Finalmente hemos presentado nuestra propuesta de estudio para desarrollar el aprendizaje cognitivo en un robot a través de la identificación de personas. En estos momentos, el robot es capaz solamente de detectar caras (y de seguirlas), pero nuestra perspectiva para el futuro es de dotarle con un sistema de reconocimiento no supervisado de caras.

Agradecimientos

El presente trabajo ha sido posible gracias al proyecto MCYT Grant TIC2003-00654 del Ministerio de Ciencia y Tecnología de España. Bogdan Raducanu es investigador del programa Ramon y Cajal, del Ministerio de Educación y Ciencia de España.

References

1. AIBO robot. <http://www.sony.net/Products/aibo/index.html>
2. Biederman, I., Mezzanotte, R.J., Rabinowitz, J.C.: Scene Perception: Detecting and Judging Objects Undergoing Relational Violations. *Cognitive Psychology*, **14** (1982) 143–177
3. Brézillon, P.: Context in Problem Solving: A Survey. *The Knowledge Engineering Review*, **14** (1999) 1–34
4. Brooks, R.A.: Intelligence without Reason. *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, Sydney, Australia (1991) 569–595
5. Brooks, R.A., Stein, L.A.: Building Brains for Bodies. *AI Memo No. 1439*, MIT (1993)
6. Bruce, V., Young, A.: *The Eye of the Beholder*. Oxford University Press (1998)

7. Fischler, M.A., Elschlager, R.A.: The Representation and Matching of Pictorial Structures. *IEEE Transactions on Computers*, **COM-22** (1973) 67-92
8. Fodor, J. A.: *The Modularity of Mind*. MIT Press, Cambridge, Massachusetts (1983)
9. Kaplan, F., Oudeyer, P-Y.: Motivational Principles for Visual Know-How Development. *Proceedings of the 3rd Epigenetic Robotics Workshop: Modeling Cognitive Development in Robotic Systems* (eds. C.G. Prince, L. Berthouze, H. Kozima, D. Bullock, G. Stojanov and C. Balkenius), Lund University Cognitive Studies, Sweden (2003) 72-80
10. Kelly, K.: *Out of Control*. Addison-Wesley, New York (1994)
11. Minsky, M.: *The Society of Mind*. Simon & Schuster Publisher, New York (1988)
12. Newell, A., Simon, H. A.: *Computer Science as Empirical Inquiry: Symbols and Search*. *Mind Design* (ed. J. Haugeland), MIT Press, Cambridge, Massachusetts (1981) 35-66
13. Nilsson, N.J.: Introduction to Machine Learning. Draft book, Stanford University (1996) Available on Internet at: <http://ai.stanford.edu/people/nilsson/mlbook.html>
14. Ormrod, J.E.: *Human Learning* (3rd Edition). Merrill Prentice Hall, Upper Sadle River, New Jersey (1999)
15. Oudeyer, P-Y., Kaplan, F.: Intelligent Adaptive Curiosity: A Source of Self-Development. *Proceedings of the 4th Epigenetic Robotic Workshop*, Genoa, Italy (2004) pp. N/A
16. Shortliffe, E.H.: *MYCIN: Computer-based Medical Consultations*. Elsevier, New York (1976)
17. Singh, P.: The Open Mind Common Sense Project. MIT Media Lab. (2002) Available on Internet at: <http://www.kurzweilai.net>
18. Torralba, A., Sinha, P.: Statistical Context Priming for Object Detection. *Proceedings of International Conference on Computer Vision*, Vancouver, Canada (2001) 763-770
19. Torralba, A., Murphy, K.P., Freeman, W.T., Rubin, M.A.: Context-based Vision System for Place and Object Recognition. *Proceedings of the International Conference on Computer Vision*, Nice, France (2003) 273-280
20. Viola, P., Jones, M.J.: Robust Real-Time Face Detection. *International Journal of Computer Vision*, **57** (2004) 137-154
21. Wada, K., Shibata, T., Saito, T., Sakamoto, K., Tanie, K.: Psychological and Social Effects of One Year Robot Assisted Activity on Elderly People at a Health Service Facility for the Aged. *Proceedings of the International of the International Conference on Robotics and Automation (ICRA)*, Barcelona, Spain (2005) 2796-2801
22. Weiser, M.: The Computer for the Twenty-First Century. *Scientific American* **265** (1991) 94-104
23. Winograd, T.: Architectures for Context. *Human Computer Interaction*, **16** (2001) 401-419
24. Yokoyama, A.: The Possibility of the Psychiatric Treatment with a Robot as an Intervention - From the Viewpoint of Animal Therapy. *Proceedings of Joint 1st International Conference on Soft Computing and Intelligent Systems and 3rd International Symposium on Advanced Intelligent Systems* (2002) paper number 23Q1-1
25. Zajdel, W., Zivkovic, Z., Kröse, B.J.A.: Keeping Track of Humans: Have I Seen This Person Before?. *Proceedings of International Conference on Robotics and Automation (ICRA)*, Barcelona, Spain (2005) 2093-2098

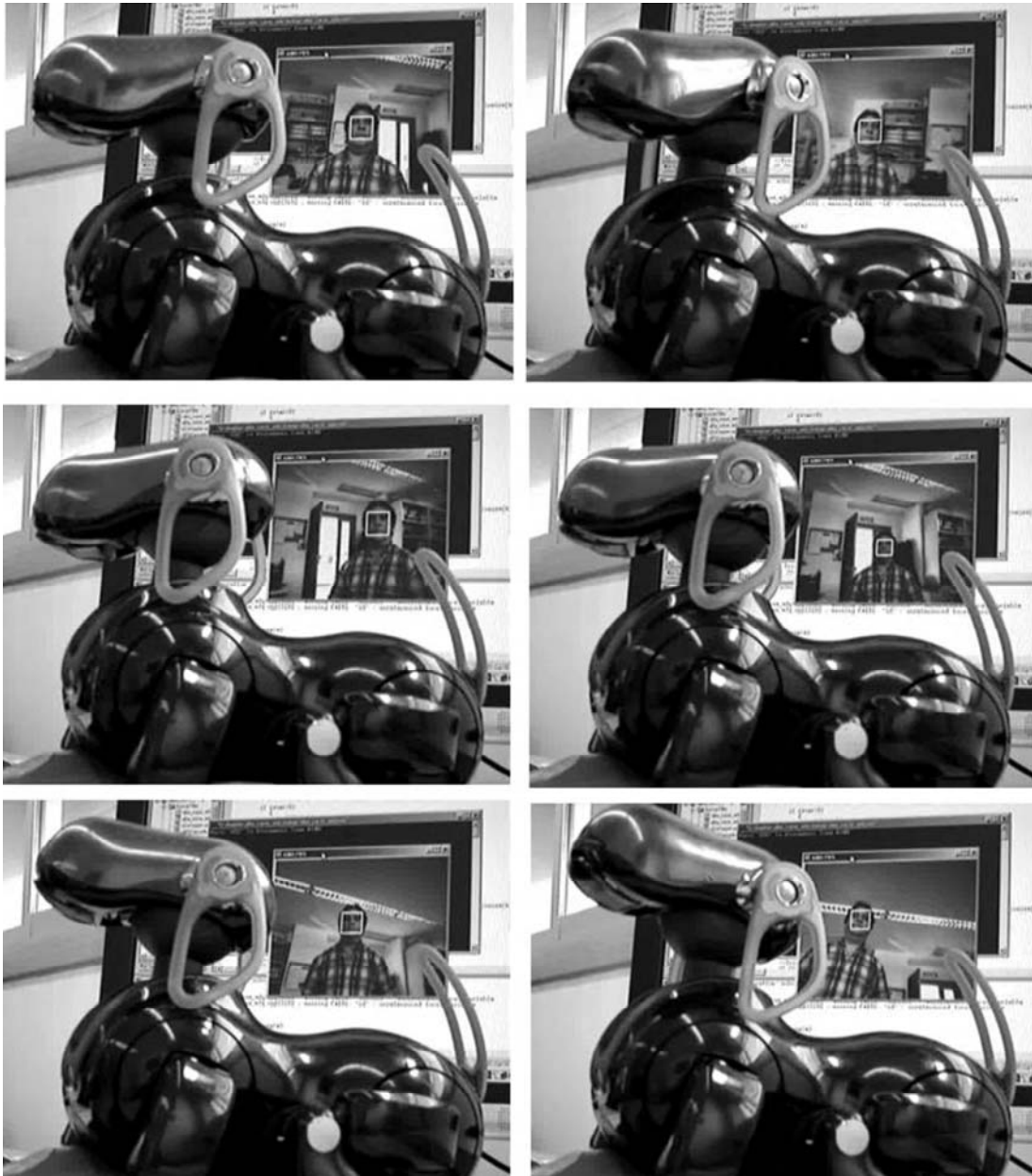


Fig. 3. Detección y seguimiento de caras en tiempo real utilizando nuestro robot AIBO. El robot ajusta la posición/orientación de la cabeza en concordancia al movimiento de la persona. Se puede apreciar la robustez del algoritmo frente a los cambios de escala e iluminación.