

ARTIFICIAL INTELLIGENCE. AN ILLUSTRATIVE OVERVIEW

Aaron Sloman

School of Computer Science

The University of Birmingham

<http://www.cs.bham.ac.uk/~axs/courses.ai.html>

Las áreas de aplicación de la Inteligencia Artificial se pueden dividir en dos, de acuerdo al contenido del estudio o de acuerdo a las herramientas y técnicas utilizadas. Se desarrollan a continuación.

(A) Contenido

Ya que los seres humanos y otros animales, y también los robots inteligentes y otros artefactos, tienen una amplia variedad de capacidades, todas ellas muy complejas y difíciles de explicar o modelar, tanto en el plano científico como ingenieril, la IA ha generado varios subcampos, tratando aspectos particulares de la inteligencia.

(B) Técnicas

Debido a que las aplicaciones de la IA son muchas y muy diversas, algunos de los subcampos se agrupan en torno a las técnicas relevantes a cada clase de problemas.

A. Subcampos basados en el contenido.

- Percepción, especialmente la visión, sino también la percepción auditiva y táctil, y, más recientemente, el gusto y el olfato. Esto se desglosa en el estudio de los diferentes tipos de procesos incluyendo la transducción física, el análisis y reconocimiento de patrones, la segmentación y "parsing" complejo de los datos sensoriales, la interpretación y el control de la atención. Este es un enorme subcampo y puede dividirse en más campos especializados de acuerdo a la modalidad sensorial, el tipo de cosas que se perciben, las formas de representación utilizadas, si la percepción está puramente dirigida por los datos o incluye procesos top-down, los mecanismos utilizados (por ejemplo neuronales o simbólicos), la arquitectura más grande que contiene el sistema sensorial, y el dominio de aplicación.
- Procesamiento de lenguaje natural, incluida la producción y la interpretación de la lengua hablada y escrita, ya sea manuscrita, impresa o electrónica en todo (por ejemplo el correo electrónico).
- Aprendizaje y desarrollo, incluidos los procesos de aprendizaje simbólico (por ejemplo la regla de inducción), la utilización de las redes neuronales (a veces descrita como sub-simbólica), el uso de algoritmos evolutivos, sistemas de auto-depuración, y diversos tipos auto-organización.
- Planificación, solución de problemas, diseño automático: dado un problema complejo y una colección de recursos, restricciones y criterios de evaluación crear una solución que cumpla con las restricciones y lo haga bien o sea óptima de acuerdo con los criterios establecidos, o si ello no es posible proponer algunas buenas alternativas.
- Variedad de razonamiento: Esto incluye estudio tanto del razonamiento informal de sentido común como de razonamiento experto especializado. El primero

incluye el estudio del razonamiento analógico, la inferencia con revocación, razonamiento basado en casos. El último incluye la lógica y razonamiento matemático, incluyendo el diseño de demostradores de teoremas y de sistemas de inferencia, ya sea con la intención de modelar diversas clases de capacidades inferenciales y matemáticas humanas, o para fines prácticos, por ejemplo, en “toolkits” de álgebra simbólica, o razonamiento en robots o sistemas de control autónomos.

- Estudio de las representaciones: la investigación de las propiedades formales de los diferentes tipos de representaciones, los mecanismos necesarios para su funcionamiento, y el tipo de tareas para las que son buenas o malas. Esto puede incluir el estudio de ontologías de diversos tipos. Sobre algunos mecanismos se afirma a veces que no utilizan ninguna representación (por ejemplo las redes neuronales), mientras que realmente son un tipo especial de representación, por ejemplo, numérica y continua, en contraposición a la estructural y discreta.
- Técnicas y mecanismos de memoria: análisis de las necesidades de los diversos tipos de memoria, incluyendo grandes almacenes de conocimiento conteniendo diversos tipos, ya sea para modelar el conocimiento humanos o para su utilización en diversos tipos de aplicaciones.
- sistemas multiagente: el estudio de los diversos tipos de comunicación (lingüística y no lingüística, explícita e implícita, intencional y no intencional), los tipos de cooperación y conflicto, reconocimiento de los planes e intenciones de otros, etc. Algunos estudios de sistemas multiagente tienen que ver con la comprensión de interacciones sociales humana, mientras que otros están preocupados con el diseño de aplicaciones que implican múltiples robots o múltiples sistemas software concurrentes. Algunos sistemas multiagente se proponen como una arquitectura para un único agente inteligente complejo.
- Mecanismos afectivos: durante los años 90 ha habido un creciente interés en el papel de la motivación y las emociones en la inteligencia. Esto se estudia a veces como un tema propio, y, a veces, como parte del estudio de arquitecturas completas para sistemas autónomos inteligentes. Una teoría general tendría que representar a una amplia variedad de estados afectivos y procesos, incluidos deseos, preferencias, antipatías, placeres, dolores, objetivos de largo plazo, intenciones, ideales, valores, actitudes, estados de ánimo, y mucho más. Uno de los debates actuales se refiere a si las emociones son necesarias para la inteligencia, o si son simplemente efectos secundarios o nuevas características de los mecanismos que se requieren para otras funciones.
- Robótica: uno de los más subcampos más antiguos de la IA. A veces estudiado con el propósito de producir nuevos tipos de máquinas útiles, y, a veces, porque diseñar completamente robots de trabajo proporciona un banco de pruebas para la integración de las teorías y técnicas de distintos subcampos de la IA, por ejemplo, percepción, aprendizaje, memoria, control motor, planificación, etc. Es decir, se trata de un contexto para explorar ideas acerca de sistemas **completos**. A veces, los diseñadores de robots intentan mostrar que ciertos tipos de mecanismos **no** son necesarios en los sistemas con un cierto tipo de inteligencia, por ejemplo, mostrando lo que pueden hacer los robots que no usan las capacidades de planificación o deliberación.

- Desarrollo de lenguajes y herramientas.
- Arquitecturas de sistemas completos. Hasta mediados de los 80 la mayor parte del trabajo en la IA se refiere a formas específicas de representación y algoritmos específicos para realizar alguna tarea. Desde entonces, se ha dado una importancia creciente a la *arquitectura* en la que muchos y diferentes mecanismos se combinan para proporcionar un sistema con muchos tipos distintos de funcionalidades, a menudo mecanismos activos simultáneamente.
- La búsqueda es otro tema que debería haber sido mencionado anteriormente, ya la búsqueda de una solución a algún problema en un espacio de posibilidades es un tema recurrente en la IA. Se han estudiado muchas formas diferentes de búsqueda, en relación a las diferentes formas de representación, diferentes dominios del problema y diferentes requisitos (por ejemplo, ¿Debe ser la solución óptima, o la meta es satisfactoria? Si no es suficientemente satisfactoria, y es muy difícil de lograr el óptimo, ¿puede bastar con hallar una solución que está garantizada de encontrarse cerca de la óptima, dentro de cierto límite?)
- Las ontologías¹ han recibido una atención considerable después de que se haya demostrado que no es suficiente especificar las formas de representación que utiliza un sistema inteligente. También es importante investigar qué tipo de cosas deberían estar representadas. Una especificación de ontología es una especificación de qué tipo de cosas tienen que existir: dos personas que comparten una ontología puede, no obstante, discrepar en cuanto a qué cosas que realmente existen en la ontología se podrían permitir, o qué leyes rigen su comportamiento. (Este tema está estrechamente vinculado a las viejas teorías filosóficas acerca de lo que existe o lo que puede existir.) El desarrollo de una ontología como resultado de la interacción con algún entorno es una clase importante de aprendizaje.

Lo anterior no pretende ser una lista completa. Hay muchos otros campos secundarios que podrían enumerarse.

B. Subcampos de aplicación de la IA

Hay un conjunto muy abierto de campos de aplicación de la IA. Los siguientes son sólo ejemplos, y no una lista completa:

- IA en la medicina, que incluye la interpretación de imágenes médicas, diagnóstico, sistemas expertos para ayudar a los médicos, la monitorización y control en las unidades de cuidados intensivos, diseño de prótesis, diseño de fármacos, sistemas tutores inteligentes para diversos aspectos de la medicina.
- IA en la robótica, que incluye la visión, el control de motores, el aprendizaje, la planificación, la comunicación lingüística, el comportamiento cooperativo.

¹ La palabra "ontología" se utiliza a menudo para referirse a lo que aquí se describe como la especificación de una ontología. Esa especificación podría ser un documento concreto en un determinado lenguaje formal. Sin embargo, la misma ontología (el mismo conjunto de las entidades, las propiedades, las relaciones, los estados, los procesos, etc.) pueden especificarse utilizando distintas notaciones, en muchos documentos diferentes. Por lo tanto, es menos confuso usar la palabra "ontología" para referirse al contenido no al vehículo, es decir, a lo que se especifica en lugar de los medios de especificar.

- IA en muchos aspectos de la ingeniería: diagnóstico de fallos, sistemas inteligentes de control, sistemas inteligentes de fabricación, ayuda inteligente al diseño, sistemas integrados de ventas, diseño, producción, mantenimiento, herramientas de configuración expertas (por ejemplo, garantizando que el personal de ventas no vendan un sistema que no funciona). AI en la ingeniería de software incluye síntesis de programas, verificación, depuración, prueba y monitorización de software.
- IA en interfaces y sistemas de "ayuda": ya que las computadoras se usan para más y más aplicaciones que implican la interacción con los seres humanos, hay cada vez más presiones para construir máquinas más fáciles de utilizar para los no expertos. Un enfoque a este problema consiste en dotar a las máquinas de más inteligencia para que puedan guiar o asesorar a los usuarios. Este tipo de aplicación puede incluir muchos de los subcampos de la sección previa.
- IA en la educación: incluye diversos tipos de sistemas tutores inteligentes y sistemas de gestión de estudiantes. Aplicaciones particulares incluyen diagnóstico de lagunas en los conocimientos del estudiante, diversos tipos de tutores de ejercicios y prácticas, marcado automático de ejercicios de programación, etc.
- IA en la gestión de la información: esto incluye el uso de la IA en la minería de datos, el rastreo web, filtrado de correo, etc.
- IA en las matemáticas: diseño de herramientas para ayudar con distintas clases de funciones matemáticas, ahora tan utilizadas que ya no se reconocen como productos de la IA.
- IA en la industria del entretenimiento: cada vez más se utiliza la IA en los juegos de ordenador y los sistemas de control y de generación de caracteres sintéticos, ya sea en la interacción a través de texto o con la generación de películas con dibujos animados o "avatares²" interactivos en mundos virtuales.
- IA en la biología: hay muchos problemas complicados en biología donde se están desarrollando sistemas informáticos más o menos inteligentes, por ejemplo, análisis de ADN, predicción de la estructura de plegado de moléculas complejas, la predicción, la elaboración de modelos de procesos biológicos, evolución, desarrollo de embriones, comportamientos de los distintos organismos.
- IA en la Ley: por ejemplo, sistemas expertos para ayudar a los abogados, o los sistemas para dar asesoramiento jurídico y ayuda a los no letrados.
- IA en la arquitectura, el diseño urbano, la gestión del tráfico: herramientas para ayudar a resolver problemas de diseño que presentan múltiples restricciones, ayudar a predecir el comportamiento de las personas en los nuevos entornos, herramientas para analizar los patrones de los fenómenos observados.

² En Internet y otras tecnologías de comunicación modernas, se denomina avatar a una representación gráfica, generalmente humana, que se asocia a un usuario para su identificación. Los avatares pueden ser fotografías o dibujos artísticos, y algunas tecnologías permiten el uso de representaciones tridimensionales.

- IA en la literatura, el arte y la música: la identificación de los autores, la modelización de los procesos de generación y el reconocimiento, las aplicaciones de enseñanza.
- IA en la detección y prevención de la delincuencia: por ejemplo, detección de falsificaciones, aprendizaje para detectar indicios de corrupción policial, software para controlar las transacciones en Internet, ayudar a planificar las operaciones de la policía, búsqueda en bases de datos policiales de evidencias de que los crímenes son cometidos por la misma persona, etc.
- IA en el comercio: Internet ha permitido que una de las áreas de mayor crecimiento en cuanto al número de aplicaciones desarrolladas sea el comercio, especialmente el comercio electrónico y el uso de agentes software de distintas clases para proporcionar, buscar, analizar o interpretar información, tomar decisiones, negociar con otros agentes, etc.
- IA en el espacio: el control a distancia de los vehículos espaciales y robots autónomos.
- IA en las actividades militares. Este puede ser el ámbito en el que se ha gastado la mayor parte de los fondos y dónde no es fácil aprender de los detalles.

Una vez más, esto no pretende ser una lista completa. Tampoco se afirma que no existan otras formas útiles de dividir la IA en posibles temas o áreas de investigación. Muchos de estos epígrafes se eligieron simplemente porque hay grupos de investigadores, conferencias, revistas, o libros que se centran en los temas dados.