Enseñanza de Inteligencia Artificial en el Nivel Introductorio: Fundamentos, Énfasis y Restricciones

Alvaro Barreiro
Dept. Computación. Fac. Informática
Universidade da Coruña
Campus de Elviña. 15071 A Coruña
email: barreiro@dc.fi.udc.es

Resumen

A partir de una consideración previa de contenidos fundamentales se propone en este trabajo el programa para un curso introductorio de Inteligencia Artificial con énfasis en Razonamiento y Acción. Se describen otros cursos alternativos y se destaca la importancia de la perspectiva integradora del diseño de agentes inteligentes en la visión global de la asignatura. Se proponen diversas aproximaciones a los trabajos de laboratorio intentando satisfacer las restricciones de créditos y formación del estudiante en los distintos paradigmas de programación.

1 Introducción

De entre las muchas definiciones de Inteligencia Artificial, la que la compara con la ciudad de Los Angeles revela alguna de las principales dificultades en la enseñanza de IA en el nivel introductorio: la diversidad del campo. Si LA son '12 suburbios en busca de una ciudad', la IA son '12 *tópicos* en busca de una materia': Búsqueda, Representación, Razonamiento, Visión, Planificación, Aprendizaje, Incertidumbre, Lenguaje Natural, Robótica, Juegos, Sistemas Expertos y Lisp [Nil95]. Podríamos ampliar la lista incluyendo el Conexionismo o campos tales como Modelado Cognitivo e IA Distribuida de difícil inclusión en la división anterior. A la dificultad de la diversidad del campo se une la de la diversidad de técnicas, puesta de manifiesto en definiciones del tipo de 'La IA estudia cómo lograr que las máquinas realicen tareas que, por el momento, son realizadas mejor por los seres humanos' [RK91] y que son consecuencia de la visión de Turing de que el comportamiento inteligente es el capaz de alcanzar la competencia propia de las personas en todas las tareas cognitivas. Esta diversidad es un hecho que se aprecia tanto en las universidades y centros de investigación como en las aplicaciones industriales que hoy en día existen. Sin embargo es una situación extraña en el curriculum de informática y puede provocar desasosiego, falta de motivación en el estudiante y dificultades para identificar las técnicas apropiadas a los problemas. Para enfrentarse a esas dificultades, un programa introductorio de IA debe considerar los siguientes aspectos.

- Los contenidos deben cubrir los fundamentos de IA completándose el programa con un énfasis determinado. Este aspecto es fundamental en la formación en un campo en continua evolución y capacita al alumno para la necesaria adquisición de nuevos conocimientos y técnicas y el manejo de nuevas herramientas en su futura actividad profesional.

- La necesidad de mostrar los contenidos desde una perspectiva integradora del campo. No son sólo motivos meramente formales sino que incluso razones cognitivas las que apoyan esta orientación pedagógica. En efecto, la aproximación integradora proporciona el procesamiento esquemático que mejora los resultados en la adquisición de la información objetivo.

Pero para que el programa finalmente tenga éxito debe satisfacer algunas restricciones. A las obvias del tiempo asignado en el plan de estudio a las lecciones y laboratorios hay que añadir las relacionadas con la formación del estudiante, especialmente importantes en la componente de programación.

En el trabajo entramos en los detalles de un curso con énfasis en Razonamiento y Acción. Se discute el papel de la programación en el curso y se propone el contenido de la componente de laboratorio. A lo largo del trabajo se describen las ventajas de la perspectiva integradora de los agentes inteligentes [RN95b] tanto para la enseñanza de contenidos teóricos como prácticos.

2 Contenidos

La propuesta de contenidos de un curso introductorio de IA debe partir de una toma de postura previa sobre lo que es fundamental. Esta es la forma de conseguir que el estudiante perciba el campo como un todo y no como una mera colección de técnicas agrupadas bajo una etiqueta. En la Tabla 1 presentamos una propuesta para un curso introductorio de 60 horas teóricas en el que se dejan 6 horas para la realización en el aula de cuestiones y problemas. Esta propuesta supone como fundamental los siguientes contenidos.

- ¿Qué es Inteligencia Artificial? En la perspectiva integradora que se adopta en la propuesta, la IA es el estudio y construcción de agentes inteligentes que perciben y actúan en el entorno [RN95a]. La IA comparte con otras ciencias (Psicología Cognitiva, Neurociencia y Lingüística) el estudio de la inteligencia pero le diferencia su objetivo aplicado de construcción de sistemas inteligentes. Sin embargo el docente debe insistir en que el objetivo aplicado no debilita el carácter científico de la asignatura (teorías, formalismos).

- Resolución de Problemas y Búsqueda. En el marco del espacio de estados el estudiante ejercita el proceso de abstracción mediante la definición formal de problemas. Los distintos tipos de problemas que surgen al estudiar la naturaleza de percepciones y acciones y la intratabilidad computacional de problemas reales son también contenidos fundamentales.

Módulo	**Tema**	**[RN95a]**	**F/E**	**Horas**
Definiciones	Definiciones	1	F	1
Agentes Inteligentes	Agentes Inteligentes	2	F	3
Resolución de problemas	Resolución de Problemas mediante Búsqueda	3	F	6
	Búsqueda Heurística	4	E	5
Lógica y Representación del Conocimiento	Lógica Proposicional	6	F	7
	Lógica de Predicados (RC)	7	F	6
	Lógica de Predicados (Inf)	9	F	8
Sistemas de Razonamiento basados en Lógica de Predicados	Programación Lógica, Sistemas de Producción y Redes Semánticas	10	E	12
Planificación	Introducción a la Planificación	11	E	6

Table 1: Curso introductorio de IA con énfasis en Razonamiento y Acción.

- Lógica, Representación del Conocimiento e Inferencia. Realmente es la Representación del Conocimiento y el Razonamiento la componente fundamental que permite que un agente tenga una representación del mundo y la utilice para tomar decisiones racionales. Esta es la visión de la Representación del Conocimiento como epistemología aplicada. Incluso admitiendo la posibilidad de distintos formalismos (lógicas, teoría de la probabilidad, conexionismo) la lógica de predicados es el más importante para la Representación del Conocimiento. Así pueden verse otros desarrollos propios de la IA como sistemas de representación del conocimiento y razonamiento basados en lógica de predicados. Debe destacarse también la importancia de la lógica como herramienta de análisis cuando se usan otros formalismos tanto en IA como en otros campos de la informática.

Una ventaja de la identificación de contenidos fundamentales es la posibilidad de completar el programa poniendo el énfasis en distintos aspectos de la IA. En el curso propuesto en la Tabla 1 se ha puesto el énfasis en Razonamiento y Acción. A los temas con contenidos fundamentales (marcados con F) se le añaden los que ponen el énfasis en esos aspectos (marcados con E). La necesidad de considerar información sobre el espacio de estados que surge de la intratabilidad computacional de la consideración de todas las secuencias de acciones se trata en el tema de Búsqueda Heurística. En el tema Programación Lógica, Sistemas de Producción y Redes Semánticas se abordan implementaciones de sistemas de representación del conocimiento y razonamiento y es una ampliación natural de los contenidos fundamentales. Este tema podría dividirse en varios pero se ha preferido adaptarlo a la división en temas que presenta el texto de Russell&Norvig

[RN95a] y que se muestra en la tercera columna de la Tabla 1. La construcción de planes de acciones con representaciones lógicas explícitas se aborda en el tema de Introducción a la Planificación. Se exponen a continuación algunas alternativas desarrolladas a partir del curso propuesto para cursos con otros énfasis. Se podría incluso pensar en cursos con orientaciones muy especializadas (Lenguaje Natural, Visión y Robótica) pero creemos que los que se muestran son los más apropiados para el nivel introductorio.

- Curso con énfasis en Razonamiento. Sustituir los temas de Búsqueda Heurística e Introducción a la Planificación por otros de Razonamiento no Monótono y Sistemas de Mantenimiento de Verdad.
- Curso con énfasis en Razonamiento e Incertidumbre. Sustituir los temas de Búsqueda Heurística e Introducción a la Planificación por otros de Teoría de la Probabilidad y Redes de Creencias.
- Curso con énfasis en Redes Neuronales y Aprendizaje . Sustituir los temas de Introducción a la Planificación y Programación Lógica, Sistemas de Producción y Redes Semánticas por temas de Redes Neuronales y Aprendizaje.

3 Perspectiva integradora

Los contenidos del curso se presentan desde la perspectiva integradora del diseño de agentes inteligentes propuesta en el texto de Russell&Norvig. Hay varias razones a favor de la enseñanza en el nivel introductorio desde una perspectiva integradora y desde la de agentes inteligentes en particular.

- La menos importante es la personal predilección por los textos escritos desde alguna perspectiva integradora. Otro ejemplo interesante es el texto de Genesereth&Nilsson [GN87] escrito desde la visión integradora de la expresividad de las lógicas para abordar los problemas de la IA.
- De manera similar a como un esquema favorece la adquisición de la información objetivo, el tema de los agentes inteligentes favorece la adquisición de conocimientos en los diversos temas del curso.
- Realmente el concepto de los agentes inteligentes está anclado en la tradición cognitiva anterior incluso a la IA (siguiendo a Johnson-Laird [JL88], Craik ignoraba la teoría de la computación, un campo recóndito y poco conocido en aquellos años):
  Si el organismo posee dentro de su cabeza un "modelo a pequeña escala" de la realidad externa y de sus posibles acciones, es capaz de ensayar varias alternativas, concluir cual es la mejor, reaccionar ante situaciones futuras antes de que aparezcan, utilizar el conocimiento de eventos pasados al tratar

con el presente y el futuro, y en cualquier caso reaccionar de una manera mas completa, segura y competente a las emergencias que se le presentan. [Cra43]

- La razón más importante es que esta aproximación permite mostrar que la IA es la tecnología apropiada para el diseño de un agente inteligente que percibe de su entorno y decide y ejecuta las mejores acciones para lograr su objetivo. De esta manera el estudiante entiende de manera natural la importancia de la identificación de objetivos, de conseguir la mejor manera de representar conocimiento y razonar con él y que un sistema de producción o un sistema de planificación son distintas maneras de representar la función que proyecta percepciones en acciones.

Además de la visión unificadora del campo el texto de Russell&Norvig presenta muchas ventajas para ser considerado el mejor en el curso introductorio de IA y de hecho ha sido adaptado en un gran número de universidades.[1] Puede verse la revisión de Nilsson [Nil96] para un análisis en más detalle pero es justo destacar al menos lo siguiente.

- Lo más importante en el nivel introductorio es explicar con claridad las ideas fundamentales. El texto de Russell&Norvig lo logra incorporando además un buen balance entre teoría y práctica. Así por ejemplo, en el tema de Búsqueda las ideas claves son la definición formal de problemas y la solución en el marco del espacio de estados. Pero estas ideas deben acompañarse del aparato teórico (complejidad y admisibilidad) y de la programación eficiente de los algoritmos de búsqueda.
- El gran número de algoritmos mostrados con seudocódigo facilita la comprensión de las ideas.[2]

4 La programación y los lenguajes de programación

En la sección 3 se ha indicado que en las lecciones la exposición de las ideas claves se acompaña no sólo de definiciones, teoremas y demostraciones sino también de cuestiones que afectan a la realización práctica de los sistemas: descripciones de los algoritmos con ayuda de seudocódigo, métodos de programación eficiente e incluso trucos de implementación. De ahí que *lecciones* sea en estos cursos un término más adecuado que el frecuentemente usado *clases teóricas*. En esta

[1] En http://www.cs.berkeley.edu/~russell/aima/adoptions.html puede obtenerse la lista de cursos que han adoptado el texto y los temas incluidos en cada curso.

[2] En http://www.cs.berkeley.edu/~russell/aima/code.html se puede conseguir el código Lisp de los algoritmos del texto.

sección vamos a referirnos al lenguaje de programación que utilizará el estudiante en las prácticas de laboratorio, a las restricciones impuestas por los planes de estudios e incluso a la posible alternativa de que la componente de laboratorio no incluya programación. Dejamos para la sección 5 algunas propuestas para los contenidos de esa componente.

Cuando los planes de estudio sean muy restrictivos en cuanto a las horas asignadas a prácticas de laboratorio y a la formación de los alumnos en programación, y en cualquier caso como actividades complementarias, pueden plantearse trabajos de laboratorio que no incluyan programación. Las aproximaciones siguientes pueden ser de utilidad.

- El uso de sistemas reales que ilustren las ideas expuestas en las lecciones. En el almacén de IA del CMU pueden encontrarse sistemas de representación del conocimiento y sistemas de planificación.

- Programas específicos para la enseñanza de IA. El ejemplo más significativo es FLAIR [IHA+94], un sistema que expone algunos temas de IA en forma de grafo conceptual y por otra parte presenta algunos algoritmos con animaciones gráficas y permitiendo la interacción del estudiante.

Cada vez hay más aplicaciones industriales de la IA en C++; el Lisp es el lenguaje de los pioneros y de la tradición pero hay que tener en cuenta que toda la informática y en particular los subcampos de IA son independientes del lenguaje [HF95]. Estoy de acuerdo con estas argumentaciones pero lo que realmente concluyen es un replanteamiento del Lisp como uno de los 12 subcampos propios de la IA en la enumeración de Nilsson, y no a la negación de su validez desde la perspectiva de la enseñanza. Precisamente su carácter tradicional ha hecho que exista abundante bibliografía especializada de Lisp en IA y accesibilidad del código fuente. La característica de lenguaje de muy alto nivel y la existencia de herramientas de desarrollo y depuración hacen que el Lisp siga siendo el lenguaje más usado en la docencia de IA.

En muchos centros se plantea el problema de que los estudiantes del curso introductorio de IA no tienen conocimientos de programación funcional y en particular de Lisp. Si se quiere incluir programación en el trabajo de laboratorio la solución suele pasar por asignar créditos de programación a los créditos de laboratorio de IA. Esta solución, si bien la única posible en muchos casos, tiene el inconveniente de desvirtuar el laboratorio de IA y el peligro de convertirlo en un curso de programación funcional. El curso cuyos contenidos hemos presentado en la Tabla 1 se completa con 30 horas de laboratorio. La inclusión de créditos de programación Lisp en esas horas supondría al menos la mitad del total de horas asignadas. Se ha propuesto como estrategia para resolver este problema la exigencia de un requisito previo de programación no imperativa (funcional en particular) para el curso introductorio de IA [KW95]. Esto se traduce en créditos de programación funcional y lógica después de los iniciales de programación imperativa. La inclusión de problemas, métodos y algoritmos propios de la IA en estos cursos aumentaría las posibilidades del laboratorio de IA.

A los estudiantes de nuestro centro se les imparte programación funcional en lenguaje Caml. Hemos decidido adaptar el laboratorio de IA a los conocimientos de Caml de los estudiantes debido a los inconvenientes arriba indicados de la inclusión de créditos de Lisp en el laboratorio de IA. Caml es un lenguaje funcional desarrollado y distribuido por el INRIA y de libre acceso.[3] Además de las características propias de un lenguaje funcional tiene un número suficiente de herramientas de desarrollo y depuración (analizador léxico y *parser*, depurador, interfaz gráfico usando Tk/Tcl, interfaz con Unix) que le convierten en una herramienta muy útil en la docencia.

5 Contenidos de las prácticas de laboratorio

Las prácticas de laboratorio basadas en programación pueden plantearse al menos de tres maneras distintas. El principio de no asignar créditos de programación al laboratorio de IA y la restricción de horas hacen que la alternativa seguida en nuestro curso sea la tercera.

- Realización de programas aislados que implementen algoritmos típicos de IA.
- Realización de proyectos o sistemas completos.
- Estudio de proyectos ya implementados y realización de modificaciones.

Describimos brevemente dos de esos proyectos y las posibles modificaciones a realizar.[4] Los proyectos tratan sobre los contenidos establecidos como fundamentales en el curso.

- El primer proyecto es un sistema que resuelve un *n-puzzle* con el algoritmo A*. En las prácticas propuestas debe adaptarse el sistema para resolver otros problemas de búqueda. Así el estudiante siente en la práctica la generalidad y utilidad de las técnicas de resolución de problemas e inventa nuevas heurísticas. Este tipo de prácticas debe incluir la constatación experimental de los resultados teóricos sobre complejidad y admisibilidad vistos en las lecciones.

[3] En http://pauillac.inria.fr/caml/FAQ/general-eng.html está La página de bienvenida a Caml.

[4] En http://www.dc.fi.udc.es/ai/~barreiro/iadocen/ialab.htm están disponibles estos proyectos.

- El segundo proyecto es una adaptación de una propuesta de S. Russell. Es un sistema de representación del conocimiento e inferencia en el que la base de conocimiento está en forma normal de Horn y los mecanismos de encadenamiento hacia delante y hacia atrás se basan en la inferencia Modus Ponens Generalizado. El manejo exhaustivo de listas y la conveniencia de analizador léxico y *parser* hacen de Caml un lenguaje apropiado para la implementación del sistema. La característica de proyecto abierto hace que puedan proponerse muchas modificaciones (mejorar las funciones de almacenamiento y recuperación en la base de conocimiento, añadir capacidades de explicación, añadir controles de inferencia recursiva, añadir capacidades de razonamiento no monótono, etc.) y que sea factible su realización en las horas asignadas de laboratorio.

6 Conclusiones

La diversidad real de la IA refuerza la necesidad de definir lo fundamental en el curso introductorio. Al mismo tiempo esa diversidad favorece los diseños alternativos de distintos cursos con distintos énfasis y aconseja plantear las lecciones desde una perspectiva integradora. La falta de formación del estudiante en los paradigmas de programación funcional y lógico suele ser la mayor restricción de la componente de laboratorio. La mejor estrategia sería la inclusión de créditos de programación funcional orientada a IA evitando la reducción de los créditos reales de laboratorio.

References

[Cra43] K. Craik. *The Nature of Explanation.* Cambridge University Press, Cambridge, 1943.

[GN87] M.R. Genesereth and N.J. Nilsson. *Logical Foundations of Artificial Intelligence.* Morgan Kaufmann, Palo Alto, CA, 1987.

[HF95] P. Hayes and K. Ford. Intellectual archeology. *ACM SIGART Bulletin*, 6(2):19–21, 1995.

[IHA+94] G. Ingargiola, N. Hoskin, R. Aiken, R. Dubey, J. Wilson, M.A. Papalaskari, M. Christensen, and R. Webster. A repository that supports teaching and cooperation in the introductory AI course. *ACM SIGCSE Bulletin*, 26(1):36–40, 1994.

[JL88] P.N. Johnson-Laird. *The Computer and the Mind: An Introduction to Cognitive Science.* Harvard University Press, Cambridge, Mass, 1988.

[KW95] D. Kumar and R. Wyatt. Undergraduate AI and its non-imperative prerequisite. *ACM SIGART Bulletin*, 6(2):11–13, 1995.

[Nil95] N.J. Nilsson. Evolutionary artificial intelligence. *ACM SIGART Bulletin*, 6(2):22–23, 1995.

[Nil96] N.J. Nilsson. Stuart Russell and Peter Norvig *artificial intelligence: a modern approach* (book review). *Artificial Intelligence*, 82:369–380, 1996.

[RK91] E. Rich and K. Knight. *Artificial Intelligence*. McGraw-Hill, New York, second edition, 1991.

[RN95a] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, Englewood Cliffs, NJ, 1995.

[RN95b] S. Russell and P. Norvig. A modern, agent-oriented approach to introductory artificial intelligence. *ACM SIGART Bulletin*, 6(2):24–26, 1995.