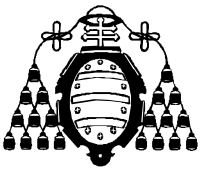


UNIVERSIDAD DE OVIEDO

DEPARTAMENTO DE INFORMÁTICA

NOTAS (EN BORRADOR) SOBRE

APRENDIZAJE AUTOMÁTICO

CENTRO:

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS INDUSTRIALES E INGENIEROS INFORMÁTICOS

TITULACIÓN:

INGENIERO INFORMÁTICO

ASIGNATURA:

ENSEÑANZA Y APRENDIZAJE ASISTIDOS POR COMPUTADOR

CURSO ACADÉMICO: 1999-2000

PROFESOR (APRENDIZAJE)
Alfredo Alguero García

SUMARIO

I. INTRODUCCIÓN AL APRENDIZAJE AUTOMÁTICO
1. Encuadre histórico.
2. Definiciones.
3. Clasificaciones.

II. APRENDIZAJE INDUCTIVO
1. Introducción y conceptos generales.
2. Inducción gramatical.
2.1. Algoritmos: IM1, KR, KS, SKR y SKC
3. Aprendizaje inductivo supervisado.
3.1. Espacios de versiones y el algoritmo de eliminación de candidatos.
3.2. Aprendizaje de reglas por recubrimiento: familia AQ, ARCO.
3.3. Aprendizaje de árboles de decisión: ID3 y C4.5.
3.4. Aprendizaje de reglas con variables: FOIL.
4. Aprendizaje inductivo no supervisado.
4.1. El clustering conceptual: la familia Cluster; Cobweb.

III. APRENDIZAJE DEDUCTIVO. SISTEMAS INTEGRADOS
1. Aprendizaje deductivo.
2. Sistemas integrados: FOCL, MOBAL

IV. APRENDIZAJE BASADO EN INSTANCIAS Y POR ANALOGÍA
1. Aprendizaje basado en instancias: método de los *k-vecinos más próximos*, NGE.
2. Razonamiento basado en casos.
3. Aprendizaje por analogía.

V. OTROS ENFOQUES DEL APRENDIZAJE
1 Enfoque conexionista: introducción a las Redes Neuronales.
1.1. El Perceptron y reglas de entrenamiento.
1.2. Redes multicapa y el algoritmo de backpropagation.
2. Enfoque evolutivo: introducción a los algoritmos genéticos
3. El aprendizaje por refuerzo.

PRÁCTICAS

Manejo de los siguientes sistemas, o análogos:

1º. INGRA. Entorno que integra diferentes algoritmos de inducción gramatical: IM1, KR, KC, SKR y SKC.
2º. AQ y ARCO. Sistemas de aprendizaje de reglas por recubrimiento.
3º. C4.5. El sistema de aprendizaje de árboles de decisión por excelencia.
4º. FOIL. Un sistema de aprendizaje de reglas con variables (lógica de primer orden).
5º. COBWEB. Un sistema de aprendizaje inductivo no supervisado.
6º. FOCL. Un sistema de revisión general de teorías.
7º. El método de los *k-vecinos más próximos*.
8º. COPYCAT. Un sistema de aprendizaje por analogía.
9º. Redes Neuronales
10º Algoritmos Genéticos

I. INTRODUCCIÓN AL APRENDIZAJE AUTOMÁTICO

1. ENCUADRE HISTÓRICO.

La necesidad de un proceso de *aprendizaje automático*, refiriéndonos con este término a programas de ordenador que sean capaces de aprender, fue apreciada en los inicios de la historia de la informática, allá por los últimos años cuarenta, época de mucho optimismo en lo que se refiere a las esperanzas depositadas en la potencia de la programación de las máquinas de calcular. Era una época de grandes expectativas sobre la capacidad de cálculo de las incipientes máquinas digitales, de su rápido desarrollo y de su evolución. Quizás esas expectativas fueron superadas, pero se infravaloró la complejidad de algunos problemas que por entonces, sin ser específicamente informáticos, intuitivamente se consideraban de menor complejidad, tanto conceptual como algorítmica. Piénsese en la traducción automática, cuyas primeras y balbucientes versiones comerciales aparecieron recientemente en el mercado creando más desengaños y críticas que satisfacciones halagos, y que por los años cincuenta y primeros sesenta era un objetivo aparentemente alcanzable a corto plazo.

Las primeras dificultades encontradas en la resolución de problemas del tipo de la traducción automática, hizo que desde el inicio se considerase como solución el desarrollo de programas que *aprendiesen*. Desde entonces, esta idea no hizo más que confirmarse como una buena estrategia -y seguramente única vía- para resolver problemas como ése y semejantes: la generación de lenguaje natural, el reconocimiento de patrones complejos, la muy recientemente planteada cuestión de dotar de sensibilidad a una máquina y en general la evolución de los ordenadores hacia la *inteligencia* -en sentido fuerte-. Ese enfoque de aquellos años, con la perspectiva que la historia proporciona, no lo consideramos visionario sino que es una salida *natural* reforzada por la superación constante en la tecnológica de los ordenadores y las impresionantes expectativas que en ese sentido se vislumbran. El hombre, guiado por su instinto de comprender la Verdad, piensa en construir máquinas inteligentes que le ayuden en esa labor y para ello tiene una vez más que seguir los pasos de la Naturaleza: es necesaria la capacidad de aprender para alcanzar la inteligencia.

En 1950 Turing esbozó los comienzos del *aprendizaje automático* con la figura del *ordenador niño* [TURING, 50]. Sería ésta una máquina dotada de la capacidad de aprendizaje con el objetivo de ir demandando y recabando la información con la que nutriese su base de conocimiento e incluso mejorase la propia tarea del aprendizaje; es decir: *aprendiese a aprender.* Turing veía este proceder como una forma eficiente de conseguir una máquina inteligente en el sentido que él mismo definió en el famoso test que lleva su nombre.

Suele citarse a Samuel como referencia obligada en el nacimiento de la IA, precisamente en la rama del *aprendizaje*. En el año 59 desarrolla un programa que aprende a jugar a las damas con bastante rapidez a base de jugar con los humanos.

Por aquellos años y parte de los sesenta los sistemas de aprendizaje estaban basados en técnicas estadísticas. Posteriormente Hunt, Feigenbaum, Simon, Winston, etc. incorporan otras estructuras -grafos, lógicas, etc.- que permiten realizar aprendizaje basado en técnicas no numéricas.

En los ochenta se amplia el interés a sistemas más allá de los que aprenden basándose en ejemplos, sin olvidarse del resurgir de las redes de neuronas como sistemas de aprendizaje. El resultado de Minsky sobre la capacidad computacional del *Perceptrón* de *Rosenblantt*, cuyo aspecto negativo fue mal generalizado [MINSKY et al., 69], había frenado las esperanzas depositadas en este paradigma. En 1982

Hopfield presentó un estudio riguroso sobre la capacidad de cómputo de las redes y describió la red que lleva su nombre, evolución de otra anterior de los años setenta. El autor de ésta última, Anderson, fue de los pocos investigadores que mantuvo su trabajo sobre las redes de neuronas a pesar del estudio de Minsky.

En la actualidad el *aprendizaje automático* acapara un interés creciente debido al nivel alcanzado por la tecnología de los ordenadores, lo que convierte en abordables los problemas de aprendizaje, normalmente de una complejidad máxima y que, aún tratándose parcialmente o con algoritmos que proporcionan soluciones aproximadas, necesitan grandes recursos informáticos.

2. DEFINICIONES.

¿Qué se entiende por aprender?. Es esta una cuestión a debatir mínimamente al comienzo de un curso sobre aprendizaje. Nosotros pretendemos simplemente reflejar la disparidad de criterios a la hora de definir un concepto tan difuso, por más que el paso de los últimos años haya ido orientando sobre los aspectos más fundamentales que se han de tener en cuenta desde el enfoque del *aprendizaje automático*.

El concepto de aprendizaje y en particular el de *aprendizaje automático* son tan imprecisos que sus definiciones generan polémica. Todos los intentos que hubo para proporcionar definiciones claras y precisas tuvieron sus correspondientes críticas, por lo que no queremos dar sensación de tener la disyuntiva resuelta eligiendo una de las definiciones como favorita. Nos limitaremos a reproducir algunas de las que alcanzaron más vigencia.

> ***Aprender denota cambios en el sistema que son adaptativos en el sentido de que le capacitan para realizar en ocasiones posteriores la misma tarea a partir de la misma población, más eficaz y eficientemente.***
>
> **Herbert Simon** (En *Why Should Machines Learn*) [MICHALSKI et al. 83, pag. 28].

Según esta definición, para que exista aprendizaje se requiere un incremento del rendimiento del sistema; sin embargo hay actividades que se pueden clasificar de aprendizaje en las que este criterio es difícil de aplicar. En estos términos critica Michalski la definición anterior que, según él, es mejorada por la siguiente

> ***Aprender supone realizar cambios útiles en nuestras mentes.***
>
> **Minsky** [MINSKY 86].

El propio Minsky considera la última definición demasiado amplia, como indica Michalski que a su vez propone como una definición mejor aún, basada en el concepto de representación del conocimiento, la siguiente:

> ***Aprender es construir o modificar representaciones de lo que ha sido experimentado.***
>
> **Michalski** [MICHALSKI et al. 86, pag. 10].

Podríamos proseguir con definiciones y sus correspondientes críticas, pero es suficiente. Aprender conlleva relacionar una multitud de conceptos no muy bien delimitada -ni precisos los conceptos- de forma que abarcarlos todos en una definición quizás sea una tarea imposible.

Una definición de *aprendizaje automático* que se puede calificar de *más formal* y que adoptamos como referencia es la siguiente:

> ***Un programa de ordenador se dice que aprende de la experiencia E con respecto a alguna clase de tareas T y rendimiento P, si su rendimiento en las tareas T, medido por P, se incrementa con la experiencia E.***
>
> **Tom M. Mitchell** [MITCHELL 97, pag. 2]

3. CLASIFICACIONES

Después de perfilar el concepto sobre el programa que se desarrolla en estas notas, se trata de hacer una breve referencia a los diferentes enfoques, paradigmas y la taxonomía del *aprendizaje automático*.

Existen tres enfoques principales en el estudio del *aprendizaje automático*:

- Uno que se refiere al análisis teórico y al desarrollo de algoritmos de aprendizaje general [NATARAJAN 91] [VALIANT].
- Un segundo que estudia el desarrollo de modelos computacionales del proceso de aprendizaje humano -modelos cognitivos-.
- El tercero que se dedica a la construcción de sistemas de aprendizaje para aplicaciones específicas.

El programa que aquí se propone está fundamentalmente centrado en el tercero de los enfoques, si bien en algún capítulo se hace referencia breve a los otros dos. El desarrollo de la disciplina que tratamos pasó por una época -años ochenta y comienzos de los noventa- en la que hubo una marcada división entre los grupos de investigación y, por tanto, de las publicaciones que seguían cada uno de los enfoques. Actualmente, hay conciencia en la comunidad estudiosa del *aprendizaje automático* de la necesidad de compaginar los tres enfoques y esa división tiende a suavizarse, pero todavía se percibe, y en cierto modo se padece, el efecto de la inercia. Particularmente nos resultan muy interesantes algunos aspectos de los tres enfoques pero el número de créditos de la asignatura para la que se propone este programa hizo que la propuesta se centrase en el tercero, que es más afín al punto de vista de la ingeniería.

Los paradigmas o modelos que se utilizan en el *aprendizaje automático* son cinco [LANGLEY, 96]: *redes de neuronas*, *aprendizaje basado en casos o instancias*, *algoritmos genéticos*, *inducción de reglas* y *aprendizaje por analogía*. Las recientes comparaciones de las que han sido objeto estos paradigmas hace que hoy se acepten como muy similares todos, al menos en cuanto a capacidad de aprendizaje se refiere. En aplicaciones específicas puede ser más eficiente uno que otro, a semejanza de lo que ocurre frecuentemente en la informática, pero las diferencias que se manejaban hace unos años están en entredicho. Cada vez está más aceptado, por ejemplo, que el aprendizaje llevado a cabo por las redes neuronales, calificado de aprendizaje subsimbólico, no es más subsimbólico que las reglas de clasificación, aunque se admite que hay aspectos que los diferencian y que hacen que cada uno de ellos tenga aplicaciones específicas. Con ánimo de aprovechar las ventajas de todos, es cada vez más frecuente el desarrollo de sistemas híbridos con varios paradigmas incorporados.

Michalski [MICHALSKI, 83] propone una taxonomía de los sistemas de aprendizaje que supone tres grupos, cada uno originado por una clasificación de los sistemas según un criterio diferente:

Por la ***estrategia utilizada en el proceso de aprendizaje***.
Por la ***representación del conocimiento extraído***.
Por el ***dominio de aplicación del sistema que aprende.***

Atendiendo a la estrategia:

i. ***Aprendizaje por rutina***.
ii. ***Aprendizaje por instrucciones***.
iii. ***Aprendizaje por analogía***.
iv. ***Aprendizaje basado en ejemplos***.
v. ***Aprendizaje por observación y descubrimiento***.

Atendiendo a la representación del conocimiento extraído.

Conocimiento extraído en forma de:

i. ***Parámetros en expresiones algebraicas.***

ii. ***Árboles de decisión.***

iii. ***Gramáticas formales.***

iv. ***Reglas de producción.***

v. ***Expresiones basadas en la lógica formal y formalismos afines.***

vi. ***Grafos y Redes.***

vii. ***Marcos (frames) y esquemas.***

viii. ***Programas de ordenador y otros procedimientos codificados.***

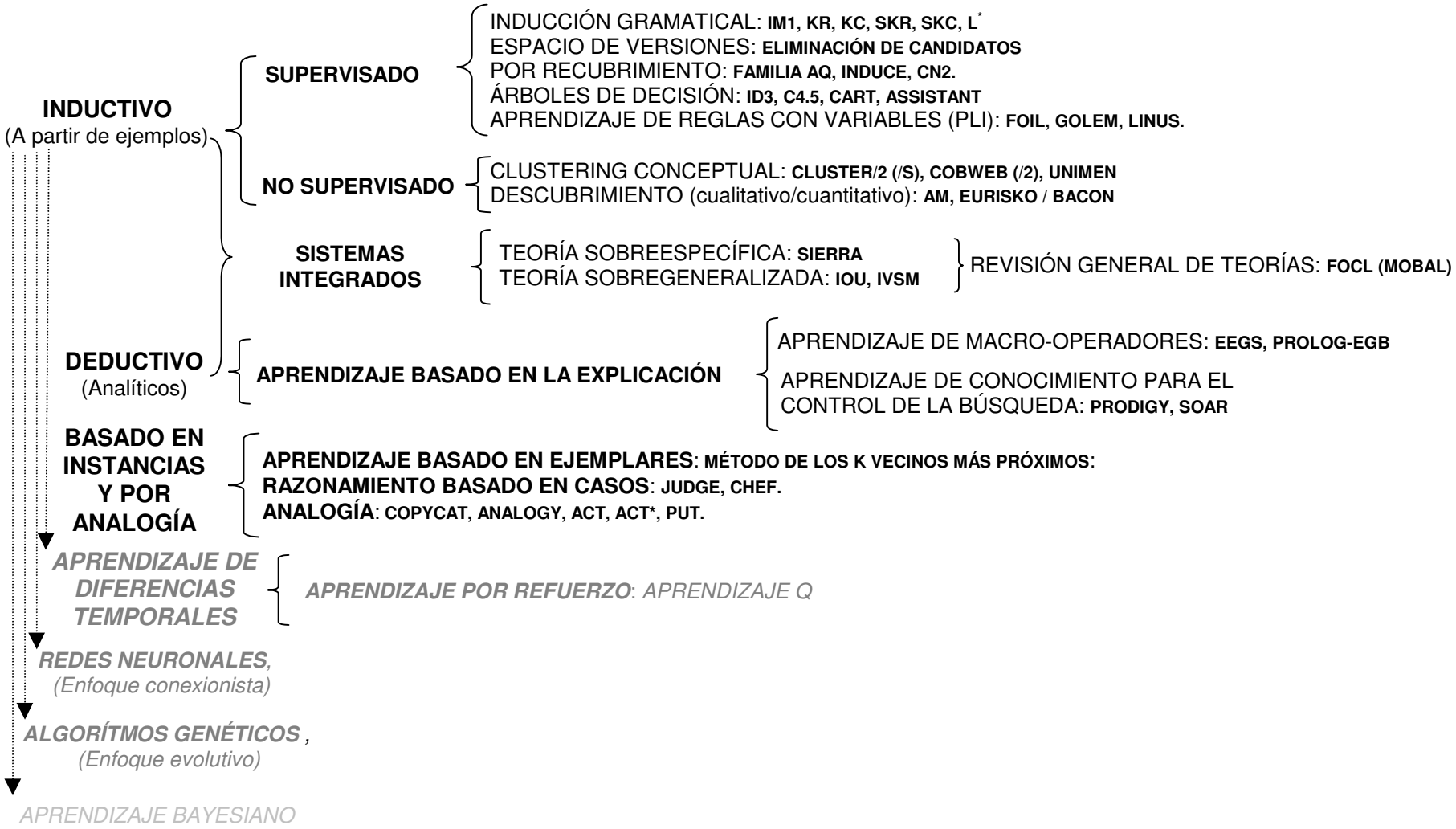
ix. ***Taxonomías.***

x. ***Representaciones múltiples.***

Atendiendo al dominio de aplicación del sistema que aprende.

No existe una clasificación cerrada por las innumerables aplicaciones posibles de estos sistemas; como ejemplos citaremos tan solo algunas de las posibles: la ***medicina***, la ***química***, la ***matemática***, la ***industria en general***, etc.

Ninguna de estas clasificaciones pueden considerarse cerradas ni de apartados estancos. La rápida evolución y la continua investigación en el tema hace que aparezcan sistemas de aprendizaje híbridos o con características que no encajan directamente en algún apartado de los anteriores, aunque normalmente suele ser posible interpretarlos como de alguno de estos tipos. Los últimos desarrollos tienden a la construcción de sistemas híbridos o multiestrategia dentro de una arquitectura que integre al aprendizaje con otras tareas como la inferencia o la representación.

En el aula se matizará brevemente cada apartado de las clasificaciones y se relacionarán éstas con la distribución de los sistemas de aprendizaje que proponemos en la **figura 1** y que sintetiza el desarrollo del programa. Prácticamente hasta el año 1996 se carecía de una metodología para le enseñanza del *aprendizaje automático*. El borrador del libro inconcluso de Nilsson [NILSSON, 96], a disposición pública en Internet para su evaluación, puede considerarse un hito bibliográfico en el sentido de proporcionar un enfoque general coherente sobre el *aprendizaje automático* aunque la publicación de la obra [MITCHELL 97] de Mitchell en 1997, prácticamente en la misma línea, lo haya relegado a un segundo plano. Creemos que la obra citada de Mitchell cumple acertadamente con el propósito con la que fue escrita: presentar un libro -podíamos de calificarlo de texto- con una visión general, conexa y suficientemente rigurosa del *aprendizaje automático*. También es digno de consideración, como complemento al enfoque de los anteriores, el libro de Briscoe y otros. [BRISCOE et al., 96] -está pendiente la publicación de un segundo volumen- que proporciona un panorámica enciclopédica de muchos de los sistemas de aprendizaje simbólico más renombrados. Estos libros mantendrán su papel de libros de referencia básicos a largo de todas las partes del programa de *Aprendizaje* que desarrollamos en estas notas, complementadas puntualmente con otros libros y artículos.

CLASIFICACIÓN DE LOS SISTEMAS DE APRENDIZAJE