

INSTITUTO NACIONAL DE MÉXICO
INSTITUTO TECNOLÓGICO IZTAPALAPA I
ORTIGOZA MARTINEZ HUITZIL
161080136
PARRA HERNANDEZ ABIEL TOMAS
INTELIGENCIA ARTIFICIAL
APUNTES

ACTIVIDADES SEMANA 3

1.1 Establecimiento de la IA y el Enfoque Simbólico

En la década de los 50 surge la primera generación de pioneros que contribuyeron al establecimiento de la inteligencia Artificial como campo de investigación, entre los que se encuentran John McCarthy y Marvin Minsky, quienes impusieron el término Inteligencia Artificial en una conferencia en 1956. Minsky se instala en el Massachusetts Institute of Technology, McCarthy emigra a la costa oeste y se incorpora a Stanford University, y Newell y Simon se establecen en Carnegie Mellon University donde lanzan distintos programas académicos y enfoques para abordar los retos planteados por la Inteligencia Artificial y se hacen pronóstico el tiempo que tomaría tener computadoras con capacidades equivalentes a la inteligencia humana.

Algunas estimaciones indicaban que este tiempo sería de unos 10 años pero debido a una subestimación de la complejidad de los problemas relacionados con la inteligencia humana, el conocimiento y el razonamiento fue un fracaso en el tiempo planteado. A lo largo de estos últimos 50 años la Inteligencia Artificial ha tenido, éxitos y fracasos, haciendo increíbles aportaciones, y nuevas promesas y predicciones sobre las capacidades de las máquinas sobre la inteligencia humana, muchas de las cuales todavía no se han cumplido.

A partir de la primera década del 2000, las tecnologías de la IA en la solución de problemas han sobrepasado por mucho las capacidades humanas como reconocimiento de voz y de imágenes en trabajos rutinarios y con una necesidad de precisión perfecta.

Los teléfonos celulares, algoritmos de planeación y búsqueda en juegos de ajedrez y Go, lenguaje natural en competencias de conocimientos con seres humanos, por mencionar solo algunas de ellas.

En 1997 la computadora Deep Blue venció al campeón mundial de ajedrez, Gary Kasparov

En 2016 el profesional surcoreano de Go, Lee Sedol fue vencido en la competencia Deep Mind Challenge por la computadora Alpha Go diseñada por Google

Aun así, esto se debe tomar con cautela ya que la comprensión profunda del lenguaje es una tarea que está todavía lejos de resolverse plenamente. En todas las

tecnologías de Inteligencia Artificial podemos encontrar dos elementos que son comunes a todas ellas:

- el empleo de conocimiento
- la utilización de métodos de razonamiento e inferencia para resolver problemas reales

Desde luego, los términos “conocimiento” y “razonamiento” evocan en principio una actividad mental humana es un hecho que las máquinas del siglo 21 tienen la capacidad de representar conocimiento y razonar, consideraremos principalmente el enfoque simbólico propuesto por Newell y Simon, y las tecnologías asociadas.

1.2. Tecnologías de Inteligencia Artificial

Tocaremos solo las IA que siguen primordialmente el enfoque simbólico basado en lógica matemática junto con probabilidad y estadística propuesto por Newell y Simon a través de la Hipótesis de los Símbolos Físicos que esencialmente establece que un sistema de símbolos operados por una máquina que sigue el modelo de la Máquina de Turing, tiene los medios necesarios para producir comportamiento inteligente general.

Es solo una hipótesis que sólo se podrá verificar su validez o su invalidez de manera experimental, pero hasta el momento, no hay evidencias que muestren que la hipótesis sea inválida.

De esta hipótesis se derivan varios enfoques el basado en lógica (los Neats) y el enfoque conexionista (los Scruffies), además del enfoque evolucionista propuesto por John Holland en 1975.

El enfoque conexionista dio origen a las Redes Neuronales ampliamente utilizadas en reconocimiento de voz, procesamiento de imágenes y en algoritmos de Aprendizaje Profundo (o Deep Learning).

El enfoque evolucionista se construye inspirándose en el paradigma de la evolución de las especies el cual es simulado a través de algoritmos genéticos y evolutivos.

En este libro nos enfocaremos solamente en el enfoque basado en Lógica Matemática como medio esencial con premisas y conclusiones representadas mediante formalismos lógicos. Los esquemas de lógica matemática a considerar incluyen la lógica proposicional y la lógica de predicados de primer orden, incluyendo lógicas descriptivas mediante ontologías y anotaciones semánticas.

Esto incluyendo el tratamiento de la incertidumbre y el razonamiento con información incompleta o incierta, el razonamiento causal, el descubrimiento de conocimiento oculto en grandes volúmenes de datos, y paquetes de software

encapsulados en la forma de agentes autónomos inteligentes, y redes de estos agentes autónomos llamados sistemas multi agente.

1.2.1 Razonamiento Lógico

El paradigma de lógica de mayor influencia en las tecnologías digitales es sin duda la lógica proposicional conocida también como lógica Booleana, que proporciona los fundamentos matemáticos para el diseño de compuertas y circuitos lógicos digitales basados en álgebra Booleana para modelar las operaciones lógicas. La lógica de primer orden o cálculo de predicados fue sistematizada antes del siglo XIX por los trabajos de Gottlob Frege y Charles Sanders Peirce. En los 60, Alan Robinson propone la regla de inferencia basada en Resolución para llevar a cabo deducciones lógicas. Programación Lógica. Prolog, qué significa Programación en Lógica, es un lenguaje de programación desarrollado en Francia por Alain Colmerauer en colaboración con el Inglés Robert Kowalski. Prolog trabaja con un subconjunto de lógica de primer orden conocido como cláusulas de Horn propuestas a su vez por Alfred Hornen 1951. una cláusula de Horn de primer orden es un problema indecidible.

El lenguaje Prolog despertó amplias expectativas para la construcción de máquinas inteligentes que inspiró al programa de computadoras de la 5o generación concebida por Japón en la década de los 80.

Otros usos y aplicaciones de la programación lógica incluyen la restricciones, la programación lógica abductiva, y el cálculo de eventos desarrollados plantea el llamado Frame Problemal usar lógica de primer orden para expresar hechos del mundo real en el modelo que un robot usa para interactuar su medio ambiente. El problema consiste en que al representar cambios en el medio ambiente, surge la necesidad de agregar una gran cantidad de hechos para representar las cosas que no cambian y que permanecen igual. Hasta la fecha, este problema se sigue investigando, y para dar respuesta a estos planteamientos se iniciaron investigaciones en otro tipo de lógicas como las lógicas de orden superior.

La ventaja de las lógicas de orden superior es su expresividad y facilidad para representar estructuras complejas de conocimiento y lenguajes. Pero esto tiene un costo, que es la complejidad exponencial generada en tiempo y espacio por los algoritmos de búsqueda y razonamiento que se hace sobre estas estructuras.

Una lógica de segundo orden es una extensión de una lógica de primer orden en la que se añaden variables que representan conjuntos, funciones y relaciones, cuantificadores que operan sobre esas variables. Con esto se incrementa el poder expresivo del lenguaje sin tener que agregar nuevos símbolos lógicos. Un tipo de lógicas de orden superior son las lógicas descriptivas utilizadas en el razonamiento semántico sobre ontologías Web.

1.2.2 Razonamiento Semántico

La WWW se creó con el objetivo de conectar personas para intercambiar y/o consumir datos. Posteriormente, estas personas podrían también producir datos. Esa gran cantidad de datos producidos carecían de significado semántico para poder ser usados en algún tipo de razonamiento o inferencia. De ahí el surgimiento de la Web Semántica. El objetivo de esta es desarrollar tecnologías para publicar datos útiles a aplicaciones computacionales, y se basa en la idea de añadir metadatos semánticos y ontológicos a sitios en la WWW que describen el contenido, el significado y la relación entre los datos. Un requerimiento es que los datos y sus metadatos se deben proporcionar por parte de los usuarios para poder procesarlas y realizar razonamientos semánticos sobre ellas. Este modelo permite incorporar contenido semántico a toda la información generada en la WWW generando una mejor organización de la información y asegurando búsquedas más precisas por significado y no por contenido textual. Finalmente, esto permite a las computadoras la gestión de conocimiento.

1.2.3 Razonamiento Probabilístico y Modelos Causales

Los problemas del mundo real se dan en ambientes de incertidumbre, información incompleta, o información inconsistente. Tradicionalmente se ha utilizado la probabilidad y la estadística como mecanismo para manejar la incertidumbre, poder interpretar eventos pasados y hacer pronóstico sobre comportamiento de eventos futuros. En los 80 's se propusieron modelos de manejo de incertidumbre como alternativa a la complejidad de tipo exponencial producida al emplear distribuciones de probabilidad multivariable, como fue el caso de los factores de certeza, pero se demostró que estos modelos resultaron ser inconsistentes. Desde entonces, las redes Bayesianas de diversos tipos se han perfeccionado para hacer frente a los problemas de incertidumbre en aplicaciones reales. Pearl también abordó el tema de la causalidad entendida como la relación entre un evento y el efecto ocasionado. Existen dos condiciones necesarias, pero no suficientes para que A sea causa de B, estas son, que A se anterior a B en el tiempo, y que A y B están relativamente cercanos en el espacio y en el tiempo. Aristóteles distinguió cuatro tipos de causas asociadas a la explicación del ser, que son la causa material, la causa formal, la causa eficiente y la causa final.

1.2.4 Razonamiento Estadístico

La estadística es la ciencia de obtener información a partir de los datos. Una estadística puede referirse a un determinado valor numérico derivado de los datos. En ese sentido este proceso se puede dividir en tres pasos: recopilar los datos, analizarlos y obtener inferencias a partir de los mismos. Esta área usa una representación de datos diferentes a las descritas anteriormente, usualmente son datos de observaciones que son almacenados en estructuras de datos tipo tabla.

1.2.5 Ciencia y Analítica de Datos

El crecimiento en las capacidades de almacenamiento de datos en dispositivos digitales, así como nuevas tecnologías de sensores y mejores anchos de banda en la transmisión de datos han permitido el almacenamiento de grandes cantidades de datos como resultado de las operaciones diarias de los negocios, los hogares, las ciudades, y en la vida cotidiana en general. Los métodos utilizados incluyen aprendizaje automático, la estadística y los sistemas de bases de datos. Incluye también aspectos de gestión de datos y de bases de datos, de procesamiento de datos, del modelo y de las consideraciones de inferencia, de métricas, teoría de la complejidad computacional, evaluación de la información descubierta, y la visualización de la información. Los pasos asociados a un proceso de minería de datos son los siguientes: Selección del conjunto de datos, tanto en lo que se refiere a las variables objetivo sobre las cuales se hará una predicción, como a las variables independientes que son las que sirven para hacer la predicción, incluyendo la muestra de datos. Análisis de las propiedades de los datos para llevar a cabo una curación o limpieza de los mismos, datos atípicos, así como datos faltantes o incompletos. Transformación de los datos de entrada para aplicar la técnica de minería de datos más apropiada a los datos y al problema. Escoger y aplicar la técnica de minería de datos para construir el modelo predictivo, de clasificación o de segmentación, dependiendo del objetivo del proyecto. Extracción de conocimiento e información útil, mediante la técnica de minería de datos seleccionada.

1.2.6 Sistemas Multiagente

Un Sistema Multiagente se compone de un conjunto de agentes autónomos con capacidad para interactuar con el medio ambiente, representar la información recibida en bases de conocimiento, y razonar con esta información para generar conocimiento, aprender, comunicarse con otros agentes, y emitir recomendaciones a otros agentes autónomos o agentes humanos.

1.3 Investigación en México

La comunidad de Inteligencia Artificial en México ha sido una de las agrupaciones académicas mejor organizadas en el país. Esto se refleja en el número de investigadores activos en el área, la cantidad de grupos de investigación en instituciones académicas, el número de estudiantes realizando estudios en el extranjero, las conferencias organizadas por entidades nacionales, los programas académicos de posgrado de IA en el país, las aplicaciones industriales de la IA, y la actividad de la Sociedad Mexicana de Inteligencia Artificial, como impulsora del desarrollo de la IA. Sin embargo, el surgimiento de la segunda guerra mundial pospuso sus planes. Wiener y Bigelow, involucrándose en proyectos de defensa, iniciaron el estudio de la retroalimentación en servomecanismos, y contactaron a Rosenblueth para que los asesorara en los problemas de retroalimentación excesiva

en el cuerpo humano, desarrollando un modelo del sistema nervioso central que podría explicar algunas de sus actividades más características en términos de procesos circulares, emergentes del sistema nervioso y dirigidos hacia los músculos, para entrar de nuevo en el sistema nervioso a través de órganos sensores. Este modelo fue publicado en 1943 en el artículo seminal «Behavior, Purpose and Teleology» en la revista Philosophy of Science. Este proponía una manera de enmarcar el concepto de propósito establecía la definición de un programa de investigación experimental y multidisciplinario. Durante los 60 's y los 70 's varios jóvenes mexicanos viajan al extranjero a realizar estudios de posgrado y al regresar se incorporan a distintas universidades del país para iniciar estudios de IA. Entre estos sobresale Adolfo Guzmán Arenas quien estudió en el MIT bajo la dirección de Marvin Minsky, uno de los fundadores de la IA. En los 80 's, se establece la Sociedad Mexicana de Inteligencia Artificial, específicamente en 1983, para formalizar e impulsar la actividad de investigación en la disciplina.

ACTIVIDADES SEMANA 4

El curso analiza qué tienen en común la búsqueda web, el reconocimiento de voz, el reconocimiento facial, la traducción automática, la conducción autónoma y la programación automática. Estos son problemas complejos en el mundo real, y el objetivo de la inteligencia artificial (IA) es utilizar herramientas matemáticas rigurosas para resolver estos problemas.

La inteligencia artificial apuesta por la construcción visual en un entorno controlado y fácil de entender. En el comportamiento de duración consideramos Inteligente, o más generalmente, el comportamiento que pensamos Posee el núcleo de la mente. Esto en sí mismo es una profunda cuestión filosófica y un intento de responder sistemáticamente. La inteligencia es un rico tema de análisis y debate. Esta pregunta sobre el marcado mental es filosófica. El desarrollo de la inteligencia artificial le ha dado una gran urgencia. El cuidadoso pensamiento filosófico sobre esta cuestión afectó el proceso de la inteligencia artificial en sí.

ACTIVIDADES SEMANA 5

Los problemas de inteligencia artificial abarcan un espectro muy amplio. Parece que tienen muy poco en común, excepto que son difíciles.

Uno de algunos resultados más difíciles y rápidos que vinieron durante las primeras tres décadas de la búsqueda de inteligencia artificial es la realización de que la inteligencia requiere conocimientos.

La técnica de inteligencia artificial es un método que explota conocimiento y debe ser representado de tal manera que:

- El conocimiento captura generalizaciones
- Entendida por personas que lo proveen

- Fácilmente modificable
- Usada para ayudar a superarse

Para crear un sistema que resuelva un problema particular (Genérico) se necesitan cuatro cosas:

1. Definir precisamente el problema
2. Analizar el problema
3. Aislar y representar el conocimiento de la tarea
4. Escoger las mejores técnicas de resolución de problemas

Programación Tradicional	Programación con IA
El programa solamente puede responder las preguntas específicas que tiene destinado a resolver	El programa con IA puede resolver preguntas genéricas que están diseñadas para resolver
Modificaciones en el programa conduce a cambios en la estructura	Modificaciones en el programa no cambia la estructura independiente de la información
Las modificaciones puede afectar negativamente el programa	El programa es rapido y facil de modificar

Las tareas de resolución de problemas pueden ser formuladas como una búsqueda en un estado de espacio.

- Un estado de espacio consiste en todos los estados de un dominio y a un conjunto de operadores que cambian un estado a otro.
- Los estados se pueden pensar mejor como nodos en un grafo conectado y los operadores como bordes.
- Algunos nodos están designados como nodos finales, y el problema se dice que es resuelto cuando un camino de un estado inicial a un estado final fue encontrado.

El conjunto de todas las posibles configuraciones de los objetos relevantes es el espacio de estados problemáticos o el espacio problemático. Esto es también llamado espacio de estado.

Espacio de estado: Se define al espacio de estado aquel que contiene todas las posibles configuraciones de los objetos relevantes.

Estados iniciales: Especifica uno o más estados dentro de ese estado como posibles soluciones a partir de las cuales la resolución de problemas pueda empezar.

Estados finales: Especifica uno o más estados que pueden ser aceptables como soluciones al problema.

Operadores: Especifica un conjunto de reglas que describen las acciones (operadores) disponibles; información que debe ser verdad, para que la acción pueda tomar lugar.

Un sistema que consta de una base de datos separada, operaciones y componentes de control representan un bloque de construcción metafórico apropiado para construir descripciones lucidas de sistemas de inteligencia artificial.

Los sistemas de inteligencia artificial muestran una separación más o menos rígida entre datos, operaciones y control.

Varias generalizaciones de los formalismos computacionales conocidos como sistemas de producción envuelve a una separación limpia de esos componentes computacionales y así parecen capturar la esencia del funcionamiento de muchos sistemas de inteligencia artificial.

Heurísticos

El estudio de los métodos y reglas para la invención y los descubrimientos. Utilizar nuestro conocimiento en el problema para considerar algunos sucesores del estado actual.

Esto haría que se podara el espacio de estados ganando velocidad pero pudiendo perder la solución.

Para hacer búsqueda heurística para trabajar, se tiene que ser posible clasificar a los hijos de un nodo.

Una función heurística toma un estado y regresa un valor numérico -- una evaluación compuesta de este estado. Entonces se elige a un hijo con el mejor resultado posible (ya sea un máximo o un mínimo).

Una función heurística puede ayudar ganando o perdiendo mucho pero encontrar la función correcta no siempre es fácil.

En una búsqueda de inteligencia artificial, la palabra heurística tiene una definición muy específica. Una heurística es una función que estima que tan cerca un estado está en la meta. La función heurística toma un estado y regresa un estimado de su distancia (costo de camino) del estado final.

La principal ganancia es la reducción del espacio de estados

- Por ejemplo el árbol completo de Gato tiene 9! ramas, si consideramos simetrías el árbol se vuelve seis veces más pequeño pero sigue siendo largo.
- Con una función heurística vamos a obtener un árbol de tan solo cuarenta estados.

ACTIVIDADES SEMANA 6

LEC 5 INFORMED SEARCH

Tipos de búsqueda heurística

- Mejor primera búsqueda
 - Algoritmo A*
- Búsqueda local
 - Hill Climbing (Montañismo)
 - Simulated annealing (Recocido simulado)

los métodos de búsqueda no informados, ya sea en amplitud o en profundidad. son métodos exhaustivos para encontrar rutas a un nodo objetivo

Estos métodos dan una solución pero no son factibles de usar porque expanden muchos nodos antes de encontrar una ruta.

los métodos de búsqueda informados usan información dependiente de tareas para ayudar a reducir la búsqueda. la información dependiente de tareas se llama información heurística y busca procedimientos usando lo que se llama métodos de búsqueda heurística.

Interesado en minimizar algunas combinaciones con el costo de la ruta y el costo de la búsqueda requerida para obtener la ruta.

Explorar métodos de búsqueda que minimicen esta combinación sobre todos los problemas que puedan ser encontrados.

La información heurística es usada para hacer búsquedas expansivas a lo largo de estos sectores de la frontera que son más prometedores.

En una búsqueda heurística cada estado es asignado a un valor heurístico que el buscador utiliza seleccionando el mejor paso siguiente.

Un heurístico es una pepita de información operativamente efectiva en cómo hacer una búsqueda directa en un espacio problemático.

Para ordenar nodos para la expansión se necesita un método para calcular la promesa de un nodo. Esto está hecho usando una función de evaluación de valor real. Las funciones de evaluación están basadas en una variedad de ideas.

Best-first search

Es la manera de combinar las ventajas de la búsqueda por profundidad y amplitud en un solo método.

- Búsqueda por profundidad es buena ya que permite una solución sin que todas las ramas tengan que expandirse
- Búsqueda por amplitud es buena porque no se queda atrapada en caminos sin salida
- Combina seguir un simple camino a la vez pero cambiando caminos cuando algún camino promete más que el camino actual.

Algoritmo A*

Minimiza el costo estimado total de la solución

Evalúa los nodos combinando $g(n)$ y $h(n)$

$g(n)$: costo de haber alcanzado n

$h(n)$: costo para llegar desde n hasta el objetivo

$f(n) = g(n) + h(n)$ costo más barato estimado de la solución a través de n

LEC 06 CONSTRAINT SATISFACTION PROBLEMS

Son problemas matemáticos definidos como un conjunto de objetos tal que su estado debe satisfacer un número de restricciones o limitaciones. Los CSP representan las entidades de un problema como una colección homogénea finita de restricciones sobre variables las que son resueltas por métodos de satisfacción de restricciones.

- En una búsqueda de estado de espacios se explora la idea de que problemas pueden ser resueltos buscando en un espacio de estados.
 - Estos estados pueden ser evaluados por una heurística de dominio específico y probada para ver si son estados finales
 - Cada estado es una caja negra accedida únicamente por rutinas de problemas específicos, la función sucesora, funciones heurísticas y pruebas finales.
 -
- Problemas de satisfacción de constantes examina problemas donde estados y pruebas finales concuerdan con la representación estándar y estructurada.
 - Se pueden definir los algoritmos de búsqueda que aprovechen la estructura de estados y usar heurísticas de propósito general en vez de específicas de problemas para generar las soluciones de problemas largos.

Un problema de satisfacción de restricción es definido como una tripla $\{X,D,C\}$ donde X es un conjunto de variables, D es el dominio de valores, y C es un conjunto de restricciones

LEC 07 SEARCHING AND/OR GRAPHS

- Representan la solución de problemas que pueden ser resueltos descomponiendolos en partes de pequeños problemas, todos de ellos tienen que ser resueltos
 - el arco AND puede apuntar a cualquier número del nodo sucesor, todos tienen que ser solucionados en el orden del arco para poner una solución
 - los arcos AND son indicados con una línea conectando todos los componentes

los grafos AND OR son definidos como hipergrafos:

- En vez de arcos conectando pares de nodos, estos son hiper arcos conectando un nodo pariente con un conjunto de nodos sucesores.
- Estos hiper arcos son llamados conectores

Para encontrar una solución en un grafo AND OR se necesita un algoritmo similar al best first search pero con la habilidad de manejar el arco AND apropiadamente.

El algoritmo debería encontrar un camino desde el nodo inicial a un conjunto de nodos representando la solución de estados.

- tal vez sea requerido alcanzar más de una solución de estado cada brazo de un arco AND debe conducir a un nodo de solución

ACTIVIDADES SEMANA 8

Logica de primer orden conocida como logica de predicados, logica cuantificacional y calculo de predicados de primer orden, es una coleccion de sistemas formales usados en matematicas, filosofia, linguistica e informatica. La logica de primer orden usa variables cuantificadas sobre objetos no logicos y permite el uso de oraciones que contienen variables de modo que en lugar de proposiciones como 'Socrates es un hombre' se pueden hacer expresiones en la forma 'Existe X tal que X es Socrates y X es un hombre' donde 'existe' es un cuantificador, mientras que x es una variable.

Esto la distingue de la logica proposicional que no usa cuantificadores ni relaciones, en este sentido, la logica proposicional es la base de la logica de primer orden.

CONCEPTUALIZACION

Se ha argumentado que la inteligencia depende del conocimiento de la entidad que tiene sobre su entorno

- Conocimiento Declarativo
 - Mucho de el conocimiento del entorno es descriptivo y puede ser expresado en una forma declarativa
- La formalizacion del conocimiento en una forma declarativa empieza con la *conceptualizacion*.
 - Incluido los objetos que presumen existir o hipoteticamente existen en el mundo y en sus interrelaciones
 - ¡Los objetos pueden ser cualquier cosa de la que queramos decir de algo!

Ejemplo

Conceptualization

Blocks World

Not all knowledge representation tasks require that we consider all the objects in the world.

Universe of Discourse – set of objects about which knowledge is expressed.

Apart from the BLOCKS; many conceptualize the TABLE on which the blocks are resting as an object as well.

In this example, there are finitely many elements in our universe of discourse. This need not always be the case.

Universe of discourse = {A, B, C, D, E}

BLOCKS WORLD scene

It is common in MATHEMATICS for example to consider the set of ALL INTEGERS as universe with infinitely many elements.

4

© Shyamara M. Thirumala, NIEL, IIT Guwahati

En este ejemplo hay un numero finito de elementos en este universo de discurso, entonces para que esta conocimiento se conceptualice se declaro como *Universe of discourse* = {A,B,C,D,E}

Muchas funciones se pueden definir. El conjunto de funciones enfatizadas a la conceptualizacion es llamado el conjunto de bases funcionales.

En el mismo ejemplo si queremos conceptualizar una parte, se puede declarar una funcion como *hat* que puede tomar la parte de un bloque que esta arriba de otro si alguno llega a existir *hat*: $\{ \langle B,A \rangle, \langle C,B \rangle, \langle E,D \rangle \}$

La relacion que es el segundo tipo de interrelacion entre objetos en este *Universo de discurso*, muchas relaciones se pueden definir, un conjunto de relaciones enfatizadas en una conceptualizacion se llama conjunto de bases relacionales.

Clear: significa que ningún bloque está encima del otro bloque

Above: alrededor de dos cuadas si y solo si está encima de la otra.

On: se mantiene si y solo si uno está inmediatamente encima del otro

Para la escena los elementos para las diferentes relaciones son:

on: $\{ \langle A,B \rangle, \langle B,C \rangle, \langle D,E \rangle \}$

above: $\{ \langle A,B \rangle, \langle B,C \rangle, \langle A,C \rangle, \langle D,E \rangle \}$

clear: $\{ \langle A,D \rangle \}$

Conceptualization

Blocks World

Many relations could be defined; The set of relations emphasized in an conceptualization is called the **relational basis set**.

clear: to mean no block is on top of the other block

A ✓

B

C

above: about two blocks if and only if one is above the other.

D

E

on: holds if and only if one is immediately above the other.

Relation – second kind of interrelationship among objects in a universe of discourse.

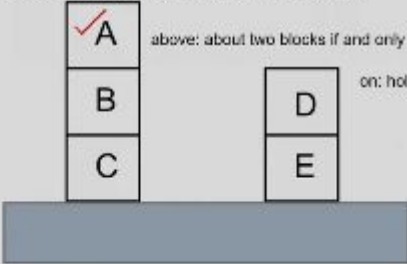
In a spatial configuration of the Blocks World, there are a number of meaningful relations.

For the scene, elements corresponding to the different relations are

on: $\{ \langle A, B \rangle, \langle B, C \rangle, \langle D, E \rangle \}$

above: $\{ \langle A, B \rangle, \langle B, C \rangle, \langle A, C \rangle, \langle D, E \rangle \}$

clear: $\{ \langle A, D \rangle \}$



BLOCKS WORLD scene

6
© Shyamshankar Mithrasara, MIT, EE Graduate

No hay una respuesta comprensible sobre que conceptualizacion es mejor que otra asi que se considera “correcta” si es consistente con el mundo.

Nosotros tenemos un conjunto de sentencias y una conceptualizacion de el mundo; nosotros asociamos simbolos usados en las sentencias con objetos, funciones y relaciones de nuestra conceptualizacion.

Con las semanticas declarativas que son la perspectiva del observador se evalua el conocimiento base con el valor verdadero de las sentencias en el mundo se puede determinar que es cierto o no

Interpretacion

Una interpretaci3n es un trazado de mapas entre elemetos de el lenguaje y elementos de la conceptualizacion.

Video 2 Administrador de proyectos

Un proyecto se define como el conjunto de las actividades que desarrolla una persona o una entidad para alcanzar un determinado objetivo, se encuentran interrelacionadas y se desarrollan de manera coordinada.

El proyecto es igual a un objetivo en común más un calendario más un costo (los dos últimos son finitos).

Normalmente en un grupo de trabajo se encuentra un **líder**, siendo este el que da dirección al proyecto, pero ¿Qué es un líder de proyecto?

El líder del proyecto es el responsable de encontrar las necesidades de los usuarios y gestionar los recursos económicos, materiales y humanos, con el fin de obtener los resultados esperados en los plazos que se hayan previsto y con la calidad necesaria, el líder debe tener ciertas habilidades para poder cumplir su función, tales como:

- Orientación al logro de objetivos
- Manejo de conflictos
- Capacidad de comunicación
- Capacidad para trabajar bajo presión

Y no solo eso, también debe de tener competencias técnicas como:

- Metodologías de desarrollo e implementación de proyectos
- Gestión de recursos humanos

Lo importante a destacar es que el líder del proyecto no es el responsable de todo lo que pase en el grupo de trabajo, todas las personas que están en el grupo deben estar comprometidos con el proyecto para que salga bien, ya de lo contrario, los cambios que puedan haber en el transcurso del proyecto o los problemas que surjan van a ser más difíciles de sobrepasar, afectando la calidad del mismo e incluso que el proyecto sea cancelado por falta de tiempo.

LEC 17 PROCEDURAL CONTROL OF REASONING

dentro de la representación y el razonamiento del conocimiento se encuentra la resolución lógica de primer orden como mecanismo de inferencia, refutación de resolución pruebas y extracción de respuestas.

es necesario recordar que la implicación en la lógica de primer orden es semidecidible, es decir, los algoritmos existen que pueden decir que si para cada oración implicada. pero no existe algoritmo que puedan decir no a toda la oración no implicada, la resolución es la refutación completa. eso es para un insatisfactorio conjunto de cláusulas alguna rama contendría la cláusula vacía, se garantiza una búsqueda en amplitud para mostrar que el conjunto de cláusulas no es satisfactorio. sin embargo, para un conjunto de cláusulas satisfactorias, la búsqueda puede o no terminar.

nuestra base de conocimientos consta de una sola fórmula, a qui esta fórmula es:

$$\forall x y z [(R(x,y) \wedge R(y,z)) \rightarrow R(x,z)]$$

mostrando la relacion R es transitiva, se podria pensar en R como una relacion que define relativo, entonces R(x,y) podria significar que X es el relativo de Y, ahora si X es una pariente de Y y Y es un pariente de Z entonces sabemos que C es una pariente de Z que es la regla en nuestra base de conocimiento.

$$(\neg R(x,y) \vee \neg R(y,z) \vee R(x,z))$$

la existencia de alguien pára todos los que no sean parientes

$$\exists x \forall y \neg R(x,y)$$

si existiera X para todo Y y no R de X,Y.
se negara la meta y se convertira en forma clausal.

$$\neg [\exists x \forall y \neg R(x,y)]$$

$$\forall x \exists y R(x,y)$$

la forma clausal aquí y la única regla disponible que puede ver que la base de conocimientos no conlleva la consulta si es negación así que una resolución debería fallar, el problema es si plantearlo como una resolución terminamos generando una secuencia infinita.

$$R(x, f(x))$$

nunca se llegará a la cláusula vacía

para la derivación de la resolución.

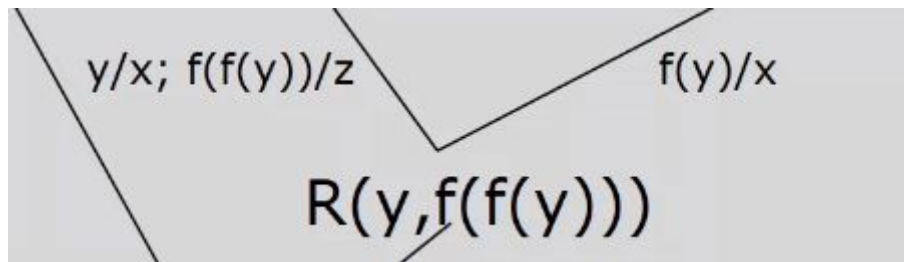
$$(\neg R(x,y) \vee \neg R(y,z) \vee R(x,z)) \quad R(x, f(x))$$

la primera cláusula que proviene de la regla en la base de conocimientos y la otra que la negación de la meta tenemos efectos de sustitución para Y y que se resuelve en la resultante que está aquí.

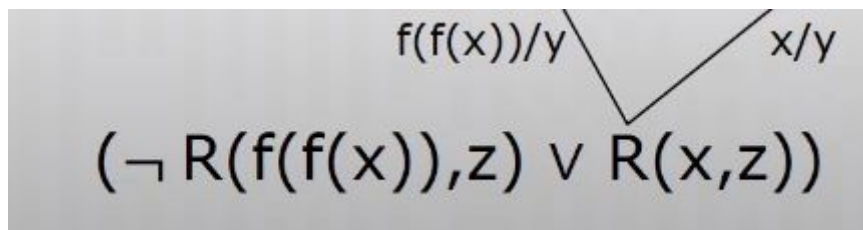
$$f(x)/y$$

$$(\neg R(f(x), z) \vee R(x, z))$$

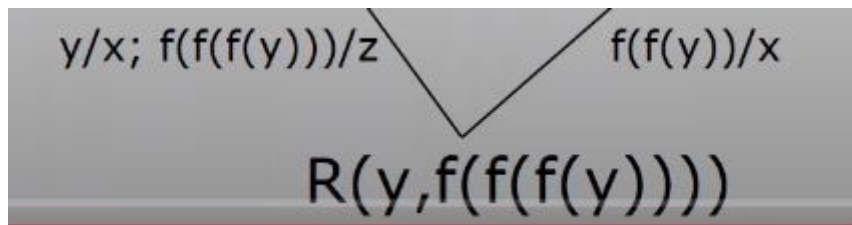
si se resuelve mas lejos, se obtiene una expresion que diga R de Y de F de F de Y una vez alli



Si se resuelve más se obtendrá una disyunción y cuando esta disyunción se resuelva tendremos una más.



cuando se resuelve la negación de la meta se obtendrá la expresión



R de Y de F de F de F de Y

si se razona en las cláusulas posteriores que se han obtenido



se puede observar que cada segundo se repite, así que no podemos utilizar una profundidad primer primer procedimiento para buscar la cláusula vacía, se puede quedar atascado.

para buscar una cláusula vacía en una ruta dada por la lógica de primer orden no hay forma de detectar si una rama continuará indefinidamente, esto se debe a:

la lógica de primer orden que uno necesita comprender es un lenguaje muy poderoso, al igual que no hay forma de detectar cuando un programa está en bucle no hay forma de detectar si una rama continuará indefinidamente.

esto es problemático desde un conocimiento perspectiva de representación, no existen procedimientos que dado un conjunto de cláusulas devuelven satisfactorio cuando las cláusulas son resolución satisfactoria sin embargo es insatisfactorio y cuando las cláusulas son satisfactorias, puede o no terminar.

por eso la resolución no es una panacea, la resolución no proporciona una eficacia general

LEC 18 REASONING UNDER UNCERTAINTY

Redes bayesianas

Las redes bayesianas son una representación gráfica de dependencias para razonamiento probabilístico en el cual los nodos y arcos representan:

Nodos variables proposicionales

Arcos: Dependencia probabilística

La variable a la que apunta el arco es dependiente (causa->efecto) de la que esta en el origen de esta

Podemos interpretar una RB de dos formas.

1. distribución de probabilidad conjunta de las variables representadas en la red representa la distribución de la probabilidad.

$$P(C,I,F,D)=P(C)P(I|C)P(F|I)P(D|I)$$

2. Base de reglas

Cada arco representa un conjunto de reglas que asocian las variables involucradas,

Por ejemplo:

Si I entonces F Dichas reglas están cuantificadas por las probabilidades respectivas.

Estructura

- La topología o estructura de la red nos da información sobre las dependencias probabilísticas entre las variables.
- La red también representa las independencias condicionales de una variable (o conjunto de variables) dadas otra variable(s).

Representación Gráfica

- Una red bayesiana representa en forma gráfica las dependencias e independencias entre variables aleatorias, en particular las independencias condicionales
- Independencia en la distribución
 - $P(X | Y,Z) = P(X | Z)$
- Independencia en el grafo

– X “separada” de Y por Z



Separación

- El conjunto de variables A es independiente del conjunto B dado el conjunto C, si no existe trayectoria entre A y B en que

1. Todos los nodos convergentes están o tienen descendientes en C
2. Todos los demás nodos están fuera de C

Tres casos básicos

Arcos divergentes->



Arcos en secuencia->



Lec 19: Bayesian Network

Puede resultar muy difícil especificar las probabilidades de los eventos atómicos; Se requiere una gran cantidad de datos a partir de los cuales se recopilan estimaciones estadísticas.

Es útil usar una notación gráfica para representar la condicional relación de independencia que se mantienen en el dominio.

En una red bayesiana :

Cada nodo representa una de las variables aleatorias del dominio.

Un arco dirigido entre dos nodos representa una relación de influencia directa entre las dos variables.

(El arco va de la variable 'causa' a la variable 'efecto').

Los nodos hermanos en el gráfico representan variables que son condicionalmente independiente dada la variable padre.

No se permite que una red bayesiana tenga ciclos.

Nodos y valores

¿Cuáles son los nodos a representar y qué valores pueden tomar, o en qué estado pueden estar? Consideraremos solo los nodos que toman valores discretos.

Los valores deben ser mutuamente excluyentes y exhaustivos.

Los nodos pueden ser nodos booleanos discretos o continuos: representan proposiciones que toman valores binarios

Ejemplo: el nodo de cáncer representa la proposición "el paciente tiene cáncer "

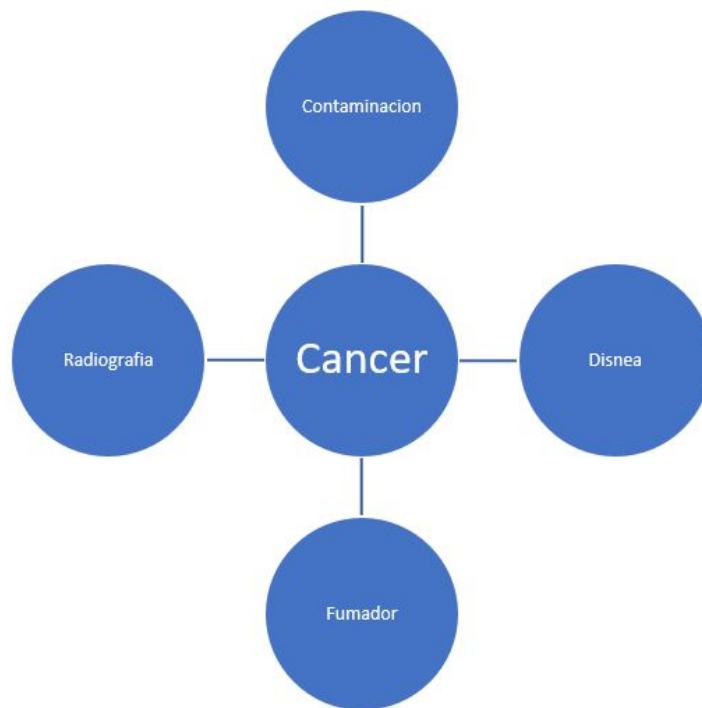
Valores ordenados

Ejemplo: nodo de contaminación con valores bajo, medio, alto

Valores integrales

Ejemplo: Edad con valores posibles 1-120

Ejemplo:



Dos nodos deben estar conectados directamente si uno afecta o causa al otro, con el arco indicando la dirección del efecto.

Tener cáncer afectará la respiración del paciente y las posibilidades de obtener un resultado radiográfico positivo. Entonces agregamos arcos de Cáncer a Disnea y Rayos X.

Terminología de estructura

Nodo es un padre de un hijo, si hay un arco desde el primero al segundo.

Para una cadena dirigida de nodos, un nodo es un antepasado de otro si aparece antes en la cadena, mientras que un nodo es un descendiente de otro nodo si viene más adelante en la cadena.

El nódulo de cáncer tiene dos padres, Contaminación y Fumador, mientras que el fumador es un antepasado de rayos X Y disnea

Rayos X es hijo de Cáncer y descendiente de Fumador y Contaminacion.

ACTIVIDADES SEMANA 13

LEC 20 DECISION NETWORK

Una red de decisión (también llamada diagrama de influencia) es una representación gráfica de un problema de decisión secuencial finito. Las redes de decisión amplían las redes de creencias para incluir variables de decisión y utilidad. Una red de decisiones extiende la red de decisiones de una sola etapa para permitir decisiones secuenciales.

En particular, una red de decisión es un gráfico acíclico dirigido (DAG) con nodos de azar, nodos de decisión y un nodo de utilidad. Esto amplía las redes de decisión de una sola etapa al permitir que tanto los nodos de azar como los de decisión sean padres de los nodos de decisión.

- Los arcos que entran en los nodos de decisión representan la información que estará disponible cuando se tome la decisión.
- Los arcos que entran en nodos aleatorios representan una dependencia probabilística.
- Los arcos que ingresan al nodo de la utilidad representan de qué depende la utilidad.

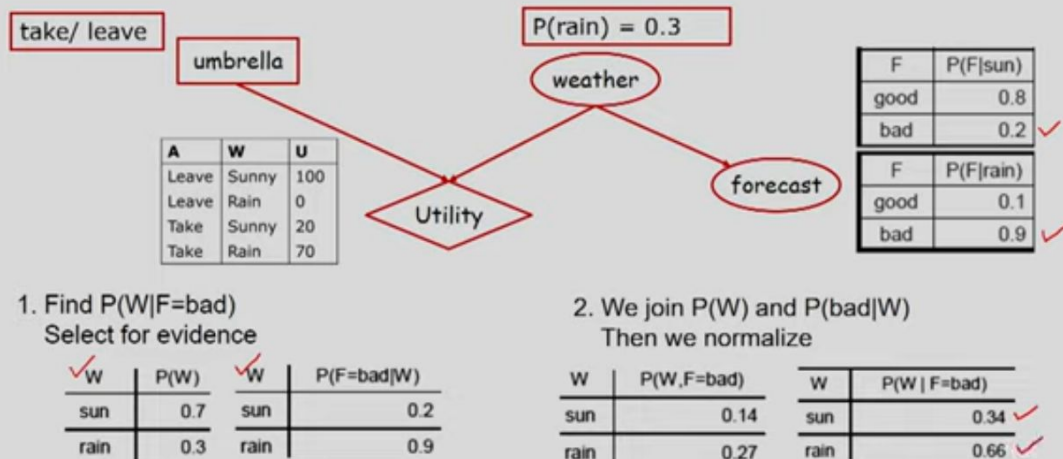
Un **agente que no olvida** es un agente cuyas decisiones están totalmente ordenadas, y el agente recuerda sus decisiones anteriores y cualquier información que estuviera disponible para una decisión anterior.

Una red de decisión sin olvidar es una red de decisión en la que los nodos de decisión están totalmente ordenados y, si el nodo de decisión D_i está antes de D_j en el orden total, entonces D_i es un padre de D_j , y cualquier padre de D_i también es un padre de D_j . Por lo tanto, cualquier información disponible para D_i está disponible para D_j , y la acción elegida para la decisión D_i es parte de la información disponible en la decisión D_j . La condición de no olvidar es suficiente para garantizar que las siguientes definiciones tengan sentido y que los siguientes algoritmos funcionen.

Uno de las partes mas importantes de tomar una decisión es conociendo que preguntas hacer, para conducir pruebas criticas dependen de dos factores:

- Que las diferentes entradas puedan marcar una gran diferencia para el curso optimo de la accion
- las probabilidades de las diferentes entradas

Decision Networks



27

© Shyamanta M. Ghosh, MIT, 2004

LEC 21 INTRODUCTION TO PLANNING

Planear es ir razonando sobre eventos futuros en orden para establecer una serie de acciones para cumplir una meta.

Un acercamiento común para planificar es representando un estado actual y determinar las series de acciones necesarias para alcanzar el estado final (o vice versa).

Los planes son creados buscando a través del espacio posibles acciones hasta que la secuencia necesaria para completar la tarea es descubierta, planear es un tipo específico de búsqueda espacial.

Planear involucra buscar a través del espacio lo cual utiliza:

- Progresión: Elige una acción donde las precondiciones se conocen hasta llegar al estado final
- Regresión: Elige una acción que marque una sub meta incompleta mientras añade precondiciones desconocidas para el grupo de sub metas. Continuar hasta que el grupo de sub metas este vacío.
 - Un enfoque hacia atrás, con orientación a objetivos, tiende a ser más eficiente

Planeadores no jerárquicos

- Método más fácil de planear
 - No hace distinción entre planes más o menos importantes
 - Se ralentiza por atender los elementos menos importantes
 - Falta de estructura entregando mal rendimiento con problemas complejos

Planeadores jerárquicos

- Toma una distinción entre las partes más importantes del plan
 - Ejemplo: Cuando compras un carro nuevo primero necesita decidir donde encuentra los fondos. No tiene sentido encontrar un buen lugar de estacionamiento en la escuela antes de tener el dinero.

Supuestos clásicos de planificación

- Información perfecta
- Efectos determinísticos

- Ejecucion instantanea
- Agente en solitario
- No preocuparse del tiempo extra, costo, recursos
- Estatico y finito

Estos supuestos fueron hechos temprano por las tareas complejas que fueron muy complejas para resolver. Estos supuestos son usados para completar tareas mas pequeñas.

LEC 22: PLAN SPACE PLANNING

Planificacion de la pila de objetivos

- Uno de los planeadores más fáciles STRIPS considera acciones razonando hacia atras; pero comprometido en acciones anticipadas.
- El algoritmo de planificación discutido es referente a la planificación de la pila de objetivos
 - Empuja sub metas y acciones a una pila; selecciona una acción solamente cuando todas las precondiciones están satisfechas.
 - Funciona en descripciones de estado donde siempre es consistente para planes crecientes, y descripciones de metas para el crecimiento del árbol de búsqueda.
- La planificación de la pila de objetivos esta incompleta, en el sentido de que puede terminar sin encontrar ningún plan.
- La planificación de la pila de objetivos rompe un conjunto de predicados de meta en submetas individuales e intenta resolverlos individualmente uno por uno.
- Este enfoque es llamado planificación lineal
 - Se refiere al hecho que las submetas se intentan y se resuelven en un orden lineal.
 - Aunque no siempre es posible.

LEC 23: PLANNING GRAPH AND GRAPHPLAN

Buscaremos soluciones en el diagrama de planificación. En la discusión, nos centraremos en la relación entre el diagrama de planificación y el espacio, y luego un brazo robótico simple que puede tomar y colocar algunos bloques, y tengo una representación mundial de los bloques. El diagrama de planificación está ubicado En su forma más simple, se considera una estructura que incorpora estados gráficos. En términos más simples, el horario es una estructura que combina los estados generados por diferentes acciones. Por lo tanto, la diferencia básica entre el espacio de estados del problema de planificación y el gráfico del plano correspondiente es que el espacio de estados se refiere a la operación del estado que genera un nuevo estado. El nuevo estado se denomina estado sucesor, que en adelante se denomina como cada acción vuelve a ser Estado, este nuevo estado se denomina estado sucesor.

ACTIVIDADES SEMANA 14

LEC 25: SEQUENTIAL DECISION PROBLEMS

En lugar de buscar soluciones en el espacio de estados o en el espacio de planificación, verá un plano gráfico con un método de planificación completamente diferente. El plano crea una estructura gráfica en la que se representan todas las posibles soluciones. A esto se le llama plan. Producido por diferentes factores. Por lo tanto, la diferencia básica entre el espacio de estados del problema de planificación y el gráfico del plan correspondiente es que el espacio de estados representa las operaciones realizadas en el estado que genera el nuevo estado, que se llama estado sucesor, y después de eso, cada acción se aplicará nuevamente. El estado de seguimiento, este es en realidad el punto de partida para una mayor exploración.

LEC 26: MAKING COMPLEX DECISIONS

en una situación de toma de decisiones no existe una opción obviamente correcta. En cambio, hay respuestas que justifican una mayor experimentación antes de comprometerte con un solo enfoque.

LEC 27: INTRODUCTION TO MACHINE LEARNING

El cálculo situacional permite que los comportamientos basados en el conocimiento infieran las consecuencias de las acciones, logrando así una representación puramente del cálculo situacional objetivos.

En su forma más simple, el diagrama de planta se considera como una estructura que combina los estados generados por diferentes operaciones.

Por lo tanto, la diferencia básica entre el espacio de estados del problema del plan y el diagrama de plan correspondiente es que el espacio de estados genera nuevos estados en el estado. El estado de la acción del estado se llama sucesor. Desde entonces, cada acción es aplicable al estado posterior, que en realidad es el punto de partida para una mayor exploración. Por otro lado, el gráfico del plan incorpora los estados producidos por diferentes acciones aplicables, por lo que tengo un conjunto de proposiciones expresadas como una sola capa. Ahora puedes ver que antes de este nivel de propuesta, habrá un conjunto de acciones que conduzcan a este nivel.

ACTIVIDADES SEMANA 15

LEC 28: LEARNING DECISION TREES

En el aprendizaje supervisado, al elemento de aprendizaje se le da el valor correcto (o aproximadamente correcto) de la función de entradas particulares, y cambia su representación de la función para tratar de calcular la información proporcionada por la retroalimentación.

Inferencia inductiva pura.

Dada una colección de ejemplos de f , devuelva una función h que se aproxime a f . La función h se llama hipótesis.

Supuesto fundamental

Supuesto: la distribución de los ejemplos de formación es idéntica a la distribución de los ejemplos de prueba (incluidos los ejemplos futuros no vistos).

- En la práctica, estas suposiciones a menudo se violan hasta cierto punto.
- Las violaciones graves claramente darán como resultado una clasificación deficiente.
- Para lograr una buena precisión sobre los datos de la prueba, los ejemplos de entrenamiento deben ser suficientemente representativos de los datos de la prueba.

árboles de decisión de aprendizaje

El aprendizaje del árbol de decisiones es una de las técnicas de clasificación más utilizadas.

- Su precisión de clasificación es competitiva con otros métodos;
- Es muy eficaz.
- Sirve como una buena introducción al área del aprendizaje inductivo.
- El aprendizaje del árbol de decisiones utiliza un árbol de decisiones (como modelo predictivo) para pasar de observaciones sobre un elemento (representado en las ramas) a conclusiones sobre el valor objetivo del elemento (representado en las hojas).

LEC 29: LINEAR REGRESSION

Variable independiente y dependiente

- Variable independiente:
 - Variable cuyo valor no cambia por efecto de otras variables y se utiliza para manipular la variable dependiente. A menudo se denota por X.
- Variable dependiente:
 - Una variable cuyo valor cambia cuando hay alguna manipulación en los valores de la variable independiente. A menudo se denota como Y.
- Las variables explicativas se denominan variables independientes y las variables a explicar se denominan variables dependientes.

Proceso de estimación

- Nuestro objetivo es conocer los parámetros del modelo que minimizan el error en las predicciones del modelo.
- Para encontrar los mejores parámetros:
 1. Defina una función de costo, o función de pérdida, que mida cuán inexactas son las predicciones de nuestro modelo.

2. Encuentre los parámetros que minimizan la pérdida, es decir, hacen que nuestro modelo sea lo más preciso posible.

Descenso gradiente: aprende los parámetros

El objetivo del descenso de gradiente es encontrar el mínimo de la función de pérdida de nuestro modelo obteniendo iterativamente una mejor y mejor aproximación de la misma.

LEC 30: SUPPORT VECTOR MACHINES

Máquina de vector de soporte lineal

- Para maximizar el ancho de la separación, es necesario minimizar w , sin dejar de respetar las restricciones con respecto a los valores en los bordes de la separación.
- Un posible enfoque para encontrar el mínimo es utilizar el método ideado por Lagrange.
- La maximización del ancho está asegurada si se minimiza lo siguiente, respetando las restricciones en los valores de los bordes.

Clasificador de margen máximo

- Formalizar la noción del mejor separador lineal.
- El mejor hiperplano es el que representa la mayor separación, o margen, entre las dos clases: se maximiza la distancia desde él hasta el punto de datos más cercano en cada lado.
- si existe tal hiperplano, se conoce como el hiperplano de margen máximo y el clasificador lineal que define se conoce como clasificador de margen máximo

Multiplicadores lagrangianos

- Manera de convertir un problema de optimización restringido en uno más fácil de resolver

Kernels

- Proyectar datos en un espacio de dimensiones superiores los hace linealmente separables.

LEC 31: UNSUPERVISED LEARNING

Aprendizaje supervisado frente a aprendizaje no supervisado

- Aprendizaje supervisado: descubre patrones en los datos que relacionan los atributos de los datos con un atributo objetivo (clase).
 - Luego, estos patrones se utilizan para predecir los valores del atributo objetivo en instancias de datos futuras.
- Aprendizaje no supervisado: los datos no tienen ningún atributo objetivo; descubrir la estructura subyacente a los datos.
 - Explore los datos para encontrar algunas estructuras intrínsecas en ellos.
 - El aprendizaje no supervisado se trata más de actividades creativas: exploración, comprensión y refinamiento.

Agrupación

- La agrupación en clústeres es una técnica para encontrar grupos de similitud en los datos, llamados grupos, es decir,
 - Agrupa las instancias de datos que son similares (cercanas) entre sí en un clúster y las instancias de datos que son muy diferentes (lejanas) entre sí en diferentes grupos.
 - En los problemas de agrupamiento, no se especifican las etiquetas de clase; sólo se conocen los vectores de

características que representan diferentes objetos / instancias / registros o situaciones.

- La agrupación a menudo se denomina una tarea de aprendizaje no supervisada ya que no se dan valores de clase que denoten una agrupación a priori de las instancias de datos, que es el caso del aprendizaje supervisado.

Métodos básicos de agrupación en clústeres

- Hay muchos algoritmos agrupados. Es difícil una categorización precisa de los métodos de agrupamiento.
- Los principales métodos de agrupación fundamentales incluyen:
 1. Agrupación en clústeres particional
 2. Agrupación jerárquica
 3. Agrupación espectral
 4. Agrupación en clústeres mediante mapas autogestionados