

## Preguntas T1 Introducción a la IA

- 1. Searle demostró que se puede hacer cosas sin ser conscientes de ello:**
  - Correcto, además, para él la conciencia es fundamental al hablar de inteligencia.
- 2. Segundo Howard Gardner existen distintos tipos de inteligencia:**
  - Cierto, entre ellas la inteligencia emocional, naturalista, cibernetica, etc.
- 3. Hofstadter postuló que una de las características de la inteligencia es responder de forma rápida a problemas o situaciones diferentes:**
  - No formas parte de las características que postulo.
- 4. Segundo Howard Gardner la inteligencia emocional:**
  - Es formada por la inteligencia intrapersonal y la interpersonal.
- 5. La inteligencia interpersonal según Howard Gardner es:**
  - La inteligencia que nos permite entender a los demás.
- 6. Segundo Howard Gardner, la inteligencia consistente es formar un modelo mental del mundo en tres dimensiones es la definición de:**
  - Inteligencia espacial.
- 7. Segundo Howard Gardner la inteligencia espacial consiste en:**
  - Formar un modelo mental del mundo en tres dimensiones.
- 8. El experimento de la “Sala china” de John Searle tiene relación con el test de Turing en:**
  - Intenta demostrar que el test de Turing no es prueba suficiente de inteligencia.
- 9. ¿Cuál de las siguientes afirmaciones acerca de la IA es incorrecta?:**
  - La IA débil a un computador actuar con un comportamiento inteligente para resolver problemas complejos siendo consciente de las acciones y decisiones que toma.
- 10. La inteligencia intrapersonal según Howard Gardner es:**
  - La inteligencia que nos permite entendernos a nosotros mismos.
- 11. Reflexionando sobre la historia de la IA, ¿Podemos afirmar que se ha conseguido alcanzar una inteligencia artificial fuerte?:**
  - No, es un objetivo distante, solo se ha iniciado el camino.
- 12. El término Inteligencia Artificial se introdujo en:**
  - La Conferencia de Dartmouth (1956).
- 13. ¿Quién diferencia varios tipos de inteligencia?:**
  - Howard Gardner.
- 14. ¿De qué es capaz una IA débil?:**
  - De resolver problemas complejos.
- 15. ¿Cuál fue el primer robot que venció a un campeón del mundo vigente de ajedrez?:**
  - Deep Blue.
- 16. Segundo Howard Gardner, la inteligencia que utilizamos cuando observamos y estudiamos la naturaleza es:**

- Inteligencia naturalista.

**17. El futuro próximo de la IA esta orientado a:**

- Crear robots capaces de abordar tareas complicadas o peligrosas para el ser humano.

**18. ¿Cuál de las siguientes seria una definición de IA?:**

- Sistemas que piensan como humanos.

**19. En la actualidad, la IA que se emplea es:**

- La IA débil.

**20. Según Howard Gardner, la inteligencia intrapersonal y la interpersonal, conforman la:**

- Inteligencia Emocional.

**21. En el campo de la IA y poder dotar de conciencia a las maquinas es cierto que:**

- Aún no se puede dotar de conciencia a una maquina ni esta previsto poder hacerlo hasta dentro de muchos años.

**22. ¿Cuál es la principal diferencia entre Turing y Searle?:**

- Uno creó un test para ver si la máquina era inteligente y el otro creó su contraejemplo.

**23. No existe ninguna diferencia entre la IA fuerte y la débil:**

- Falso.

**24. En que década se produjo la Etapa de Expansión de la IA:**

- En los 80.

**25. ¿En que conferencia y en que año se definió el campo de la inteligencia Artificial?:**

- En la conferencia de Dartmouth en el año 1956.

**26. Según Gardner, la inteligencia que nos permite entender a los demás es:**

- Inteligencia interpersonal.

**27. Según Howard Gardner, la inteligencia emocional de que otras inteligencia consta?:**

- Inteligencia intrapersonal e interpersonal.

**28. En cuanto al futuro próximo de la IA, podemos concluir que se orienta hacia:**

- Abordar aquellas tareas que, ya sea por lo incomodo, peligroso o complicado, conviene delegarlas en sistemas inteligentes artificiales.

**29. Según Howard Gardner, la inteligencia que esta formada por la inteligencia intrapersonal y la interpersonal y juntas determinan nuestra capacidad de dirigir nuestra propia vida de manera satisfactoria, se denomina:**

- Ninguna de las anteriores.

**30. La diferencia entre la IA débil y la IA fuerte es:**

- LA IA fuerte es la creada para pensar y ser consciente, y la IA débil solamente se comporta de manera inteligente.

**31. ¿En que consistía el test de Alan Turing?**

- Demuestra la existencia de inteligencia en una máquina.

**32. ¿Cuál de las siguientes afirmaciones es correcta?:**

- ~ Las IAs fuertes se relacionan con el test de Turing
- ~ Las IAs débiles se relacionan con el test de Searle
- Ninguna de las anteriores.~

**33. ¿Superar el test de Turing es condición suficiente para determinar que una máquina es inteligente?:**

- No, ya que la máquina puede estar siguiendo reglas sin que haya inteligencia.

**34. Si un robot puede resolver ciertos problemas complejos, podemos afirmar con total seguridad que posee una Inteligencia...:**

- Débil, puesto que el ordenador no tiene conciencia como un ser humano, simplemente resuelve problemas modelando su inteligencia.

**35. ¿Para Hofstadter cuál sería la definición de Inteligencia Interpersonal?:**

- ~ Es la que nos permite entendernos nosotros mismos.
- ~ Es la que nos permite entender a los demás.
- Ninguna de las anteriores.

**36. El test de Searle en su contraejemplo del test de Turing nos sirve para apreciar:**

- La IA fuerte.

**37. Si un computador se comporta de manera inteligente, se puede afirmar que es realmente inteligente de la misma forma en la que lo es un ser humano. Elige la opción correcta:**

- ~No, según Searle.
- ~Si, según Alan Turing.
- Ambas respuestas son correctas.

**38. ¿Cuales de los siguientes sistemas inteligentes pertenecen a los llamados sistemas constructivistas?:**

- ~Los sistemas que actúan como humanos.
- ~Los sistemas que actúan racionalmente.
- Ambas respuestas son correctas.

**39. El término Inteligencia Artificial se introdujo en...:**

- La Conferencia de Dartmouth (1956).

**40. Según Minsky, la Inteligencia Artificial es...:**

- ~ El estudio de las facultades mentales mediante el uso de modelos computacionales.
- ~ El estudio de los cálculos que hacen posible percibir, razonar y actuar.

- Ninguna de las anteriores es correcta.

**41. ¿Cuáles son algunas de las principales áreas de aplicación de la IA en la actualidad?:**

- ~ Juegos, Minería de datos y Predicción financiera.
- ~ Problemas de percepción, Robótica y Mundos virtuales.
- Las dos opciones anteriores son correctas.

**42. Si una persona A, le gana en una partida de ajedrez a otra persona B, ¿significa que la persona A es mas inteligente que la persona B ?:**

- No, aunque le ganes al ajedrez a otra persona no implica que seas mas inteligente.

**43. ¿Como se relacionan los siguientes términos? "Pensar como", "actuar como", "IA fuerte", "IA débil", "tener conciencia" "no tener conciencia" "Test de Touring" y "La sala china":**

- Hay dos grupos, el primero es: "Pensar como", "IA fuerte", "tener conciencia", "La sala china" y el segundo grupo es: "actuar como", "IA débil", "no tener conciencia" "Test de Touring".

**44. ¿Cuál de las siguientes afirmaciones es correcta?:**

- Para Searle, la inteligencia artificial se basa en la capacidad de una maquina para ser consciente como un humano, característica principal de una IA fuerte.

**45. ¿Qué definición de Inteligencia Artificial (IA) se adapta mejor al criterio de la asignatura?:**

- "El arte de construir máquinas capaces de hacer cosas que requerirían inteligencia si las hicieran los seres humanos". (Minsky, 1986).

**46. Según Hofstadter, la inteligencia es la habilidad para:**

- ~ Encontrar similitudes entre situaciones diferentes.
- ~ Dar sentido a mensajes ambiguos o contradictorios.
- Todas las respuestas son correctas.

**47. La inteligencia Intrapersonal....:**

- Es la que nos permite entendernos a nosotros mismos.

**48. Según los tipos de inteligencia de Howard Gardner, ¿Cuál de ellos consiste en entendernos a nosotros mismos?:**

- Inteligencia intrapersonal.

**49. ¿Que es la inteligencia?:**

- Es el conjunto de habilidades desarrolladas por el hombre para recibir información, analizarla, comprenderla, almacenarla y saberla aplicar en el futuro para la resolución de problemas.

**50. Alcanzar la conciencia se plantea como un objetivo de:**

- IA Fuerte.

**51. ¿Superar el test de Alan Turing nos garantiza que la maquina tendrá una IA fuerte?**

- No, Searle con “la sala china” demuestra que el test de Turing no es suficiente ya que una máquina puede realizar una acción sin saber que hace y por qué lo hace.

**52. Asistentes de voz como Siri (Apple) o Voz S (Samsung), que son capaces de responder a cualquier pregunta en nuestra en lenguaje natural, concluimos que pasan correctamente el test de Turing:**

- Verdadero, ya que según dicho test no es posible distinguir entre máquina y ser humano.

**53. ¿Qué nos quiere dar a entender Searle con el experimento de la sala china?:**

- Que el Test de Turing no tenía razón, porque si un computador realiza un trabajo no significa que sea consciente de lo que hace.

**54. La definición del campo de la I.A se fijó en:**

- La conferencia de Dartmouth (1956).

**55. Según Howard Gardner la inteligencia corporal-kinestésica es la inteligencia de los:**

- Cirujanos, artesanos, deportistas y bailarines.

**56. ¿Cuál de las siguientes listas representa aspectos de la inteligencia?:**

- La memoria, la resolución de problemas y la creatividad.

**57. Si un robot es capaz de resolver problemas complejos y además posee inteligencia emocional. ¿Cuál de las siguientes afirmaciones es la más correcta?:**

- Pasaría el test de Turing.

**58. Sobre el comportamiento inteligente:**

- Las máquinas pueden manifestar un comportamiento inteligente para resolver un determinado problema.

**59. Searle mediante el contraejemplo del test de Turing afirma que:**

- Actualmente, una máquina no puede ser del todo inteligente. (no es suficiente el test de Turing).

**60. Acerca de Deep Blue se puede decir que:**

- ~ Es un ejemplo de la IA fuerte, ya que venció el gran jugador de ajedrez Garry Kasparov. Así, demostrando ser más inteligente que un ser humano.
- ~ No puede ser considerado como una IA porque no es consciente de sus movimientos.
- Ninguna de las anteriores es correcta.

**61. Un computador es considerado inteligente si:**

- ~Puede acumular y utilizar el conocimiento.
- ~Son capaces de ver, entender e interpretar correctamente los estímulos sensoriales.
- Ambas respuestas son correctas.

**62. ¿Cuál de las siguientes definiciones no se corresponde con la Inteligencia Intrapersonal?:**

- Aquella que nos permite entender a los demás.

**63. ¿Como podemos dotar a una maquina de una IA fuerte?:**

- Proporcionándole la capacidad de procesamiento e inteligencia necesarios hasta el punto de alcanzar el pensamiento consciente.

**64. En el famoso experimento mental de John Searl "La habitación china", popularizado por Roger Penrose, una computadora:**

- Maneja información sintáctica; respuestas en base a una serie de reglas predefinidas sin conciencia alguna de la propia acción.

**65. Según Howard Garden, la inteligencia cibernetica es:**

- LA que desarrollan las personas estudiando y aprovechando la ciencia que se ocupa de los sistemas de control y telecomunicaciones.

**66. ¿Cual de las siguientes afirmaciones es correcta?:**

- ~ La Inteligencia artificial fuerte en un computador se caracteriza por la capacidad de este de resolver problemas complejos pero sin que este sea consciente de ello.
- ~ Cuando un computador se comporta de manera inteligente se puede afirmar que este posee una Inteligencia artificial fuerte, tal y como la de un ser humano.
- Ninguna de las anteriores.

**67. Según Howard Gardner, los que tienen Inteligencia Lingüística son:**

- Los escritores.

**68. ¿En que estado se encuentra actualmente la IA?:**

- No disponemos de una IA completa, lo que nos obliga a afrontar problemas concretos.

**69. Según la clasificación de Howard Gardner cuantos tipos de inteligencia existen:**

- 8.

**70. ¿Es posible decidir si una persona es más inteligente que otra?:**

- Depende del tipo de inteligencia.

**71. ¿Que tipo de inteligencias crees que debe destacar en un piloto de formula 1 o motociclismo?:**

- Inteligencia espacial.

**72. La conciencia de las maquinas es un objetivo a:**

- Largo plazo.

**73. Según Hofstadter, cuál de las siguientes afirmaciones no corresponden a la inteligencia:**

- Responder rígidamente a diferentes situaciones.

**74. Una IA fuerte es aquella que se caracteriza por:**

- Conjunto de procesamiento y capacidades en la cual la máquina tiene conciencia de las actuaciones humanas que está realizando o pensando.

**75. Según Howard Gardner, la inteligencia emocional es:**

- La formada por la inteligencia intrapersonal y la interpersonal.

**76. Según Searle y el ejemplo de la habitación china:**

- La conciencia es un requisito fundamental para crear una IA fuerte.

**77. Según Hofstadter, si eres inteligente eres capaz de:**

- Encontrar similitudes entre situaciones diferentes y generar nuevos conceptos e ideas innovadoras.

**78. Señala la respuesta correcta:**

- ~ El ordenado Deep Blue, ordenador que ganó al entonces campeón del mundo de ajedrez, demostró tener una buena inteligencia cibernetica.
- = En el caso de haber existido, Hal 9000 estaba dotado de una inteligencia similar a la humana ya que era consciente de todos sus pensamientos.
- ~ En un futuro muy cercano, con los avances que se ha conseguido hasta el momento, disfrutaremos de sistemas con IA fuerte.

**79. Según Howard Gardner la Inteligencia Intrapersonal es:**

- La capacidad de entendernos nosotros mismos.

**80. Según Howard Gardner ¿Qué es la inteligencia kinestésica?:**

- Es la capacidad de utilizar el propio cuerpo para realizar actividades o resolver problemas.

**81. Cuál de las siguientes definiciones se refiere a los “Sistemas que piensan racionalmente”:**

- "El estudio de los cálculos que hacen posible percibir, razonar y actuar". (Winston, 1992).

**82. ¿Qué demostró el experimento de la sala china con respecto al Test de Turing?:**

- Que el test no era infalible, porque no diferenciaba entre inteligencia fuerte y débil.

**83. Durante los años dorados de la historia de la IA nació el siguiente lenguaje de alto nivel (que aún se usa en la actualidad) en los laboratorios del MIT, siendo el lenguaje dominante en la IA en los años venideros:**

- LISP.

**84. De las siguientes acciones, cuál de ellas podría pertenecer a sistemas que actúasen como humanos:**

- Construir máquinas capaces de hacer cosas que requerirían inteligencia.

**85. ¿ La victoria de Deep Blue sobre Kasparov al ajedrez demuestra que Deep Blue es una IA fuerte ?:**

- No, Deep Blue no era consciente de lo que hacía.

**86. Se considera que la computadora HAL es inteligente:**

- Segundo Searle porque tiene conciencia de sí misma.

**87. La inteligencia espacial según Howard Gardner es un tipo de inteligencia característica de:**

- Marineros.

**88. ¿Cuál de las siguientes afirmaciones acerca de la inteligencia es correcta?:**

- La memoria, el aprendizaje y la creatividad son algunos aspectos de la inteligencia.

**89. Suponiendo que se crea un robot cuya función es realizar acciones humanas (sonreír, llorar, hacer muecas, ...) ¿Qué clase de I.A. tendría? :**

- I.A. Débil.

**90. En el futuro de la IA, entre otras tareas, se pretende abordar...:**

- Tareas incómodas, peligrosas o complicadas para el ser humano que convendrá delegar.

**91. Que tipo de inteligencia según Howard Gardner determina nuestra capacidad de dirigir nuestra vida de manera satisfactoria?:**

- La combinación de ambas.

**92. La IA fuerte:**

- Con la suficiente capacidad de procesamiento e inteligencia, se puede crear un ordenador que pueda pensar y ser consciente de la misma forma que lo hacen los seres humanos.

**93. Que es el efecto Flynn?:**

- La subida continua, año por año, de las puntuaciones de cociente intelectual, un efecto visto en la mayor parte del mundo.

**94. ¿Pueden existir personas que al dialogar sin saberlo con un chatbot (programa que simula mantener una conversación con una persona) no se percatan de hablar con un programa?**

- Verdadero. Utilizando un lenguaje muy simple, de modo tal que se cumple la prueba de Turing como cuando se formuló.

**95. ¿Cuál fue uno de los principales aportes al campo de la IA por parte de Howard Gardner?:**

- Propuso el modelo de la teoría de las inteligencias múltiples.

**96. Suponemos que aplicamos el test de Alan Turing (1950) sobre una máquina cualquiera. Diremos:**

- La máquina es inteligente si el juez que aplica la prueba no consigue discernir entre la máquina y el humano.

**97. ¿Cuál es el modelo más adecuado en IA?:**

- ~Modelo evolutivo.
- ~Modelo basado en abordar todos los problemas en una única solución.
- =Ninguna de las anteriores.

**98. Richard Herrnstein y Charles Murray llamaron efecto Flynn a:**

- La existencia de una subida continua en las puntuaciones de cociente intelectual en sucesivas generaciones.

**99. ¿En qué está orientado el futuro de la IA cuando existe la posibilidad de apoyar o delegar las tareas en sistemas inteligentes artificiales?:**

- Tareas incómodas, peligrosas o complicadas

**100. Según Searle, ¿por qué una máquina no es inteligente sólo por comportarse de manera inteligente?**

- Porque no es consciente de lo que está haciendo.

**101. ¿Qué afirmación es correcta y según quién?:**

- John Searle, que haya pasado el test X, no asegura que el programa sea inteligente.

**102. Según Hofstadter, la inteligencia es la habilidad para...:**

- Saber aprovechar circunstancias fortuitas.

**103. En la historia de la Inteligencia Artificial, ¿En qué época suceden años de crítica y madurez?:**

- Los difíciles años 70.

## Preguntas T2 Estrategias de Búsqueda

1. En el problema de las jarras de agua, siendo las reglas de producción:

1	$(x, y)$ si $x \leq 4$	$\Rightarrow (4, y)$
2	$(x, y)$ si $y \leq 3$	$\Rightarrow (x, 3)$
3	$(x, y)$ si $x \geq 0$	$\Rightarrow (x-d_1, y)$
4	$(x, y)$ si $y \geq 0$	$\Rightarrow (x, y-d_1)$
5	$(x, y)$ si $x \geq 0$	$\Rightarrow (0, y)$
6	$(x, y)$ si $y \geq 0$	$\Rightarrow (x, 0)$
7	$(x, y)$ si $x+y \geq 4$ e $y > 0$	$\Rightarrow (4, y-(4-x))$
8	$(x, y)$ si $x+y \geq 3$ e $x > 0$	$\Rightarrow (x-(3-y), 3)$
9	$(x, y)$ si $x+y \leq 4$ e $y > 0$	$\Rightarrow (x+y, 0)$
	$(x, y)$ si $x+y \leq 3$ e $x > 0$	$\Rightarrow (0, x+y)$

Dada esta secuencia de hechos y acciones:

Jarra de 4 l.	Jarra de 3 l.	Regla a aplicar
0	0	2
0	3	9
3	0	2
3	3	6
4	2	5
0	2	9
2	0	Solución

1. La secuencia de reglas obtiene la solución, es cierto:

- Falso, porque la regla 6 está mal aplicada

2. El algoritmo de ponderación dinámica es únicamente aplicable...

- a problemas donde conocemos la profundidad.

3. Una función heurística  $h(n)$  se dice que es admisible si cumple:

-  $h(n) \leq h^*(n) \forall n$

4. En el algoritmo A\* visto en clase, y para un problema en el que no podemos movernos en diagonal, la heurística óptima utiliza la distancia:

- De Manhattan.

5. ¿En qué estrategia tentativa NO INFORMADA podemos encontrar una variación del backtracking?

- Búsqueda en profundidad.

**6. El objetivo de la téctica de ajustes de pesos es:**

- definir una función  $f()$  ponderada,  $f_w()$ , como alternativa a la utilizada en A\*

**7. En la forma de proceder del algoritmo A\*:**

- Se expanden los nodos adyacentes al que se está explorando y se explora el de menor coste.

**8. Un sistema de producción es una terna entre la Base de Hechos, las Reglas de Producción y la Estrategia de Control. Señala la definición correcta:**

- La BH es el conjunto de representaciones de uno o más estados por los que atraviesa el problema.

**9. Las estrategias a considerar dentro de las estrategias de búsqueda básica son...**

- Irrevocables y tentativas.

**10. Al relajar la restricción de optimalidad (marca la respuesta incorrecta):**

- Siempre obtenemos la mejor solución.

**11. La EC(Estrategia de control) { ~ Determina el conjunto de reglas aplicables mediante filtrado y resuelve conflictos entre varias mediante pattern-matching.**

- Determina el conjunto de reglas aplicables mediante pattern-matching y resuelve conflictos entre varias mediante filtrado.

**12. La bondad de una solución es:**

- Si aceptamos soluciones aproximadas para el problema es el margen de error.

**13. Una buena heurística para el problema del 8-puzzle es... :**

- Sumar las distancias de las piezas a sus posiciones en el objetivo

**14. Respecto a las estrategias desinformadas (No informadas):**

- "son ciegas en el sentido de que el orden en el cual la búsqueda progresá no depende de la naturaleza de la solución que buscamos".

**15. Sobre los árboles de búsqueda en general. ¿A que tipo de dato responde su estructura?**

- Grafo dirigido

**16. Si tenemos un algoritmo A\* para resolver el problema de camino mínimo en el cual  $h(n)$  devolviese siempre 0, estaríamos ante un tipo de búsqueda**

- En anchura

**17. En un sistema de producción las reglas de producción (RP):**

- Cada regla tiene dos partes, Precondiciones y Postcondiciones

**18. ¿Cuál de las siguientes estrategias desinformadas es una variación del backtracking?**

- Búsqueda en profundidad

**19. Tratándose de una búsqueda A\* (búsqueda óptima) decimos que  $h^*(n)$  es:**

- Coste del camino de coste mínimo de todos los caminos desde el nodo n a cualquier estado solución  $t_j$ .

**20. Según la formalización de los problemas de búsqueda, realizado por POST en 1943, un sistema de producción se basa en una terna donde existe:**

- Una base de hechos, unas reglas de producción y una estrategia de control.

**21. ¿Qué es una lista focal?:**

- Es una sublistas de ListaFrontera que contiene solo nodos con  $f(n)$  menor al mejor valor de los  $f(n)$  de listaFrontera por un factor.

**22. El algoritmo de búsqueda A\* constituye una:**

- Estrategia tentativa informada.

**23. Las reglas de transformación**

- Son las reglas que nos permiten pasar de un estado a otro.

**24. Para resolver un problema como una búsqueda en un espacio de datos hay que:**

- Especificar las reglas de transformación.

**25. Cuando hablamos de la técnica de ponderación dinámica, sucede que:**

- Al principio, en los niveles iniciales, las heurísticas pierden admisibilidad.

**26. Para el algoritmo de búsqueda A\* (Búsqueda óptima) decimos que  $f^*(n)$  es:**

- Coste del camino de coste mínimo desde el nodo inicial hasta un nodo solución condicionado a pasar por n.

**27. Al elegir la estrategia de control, hay que tener en cuenta que la solución..:**

- Las dos anteriores son correctas.

**28. Una búsqueda en profundidad es:**

- Una estrategia tentativa.

**29. Las estrategias irrevocables son aquellas que...**

- no se permite la vuelta atrás.

**30. En lo referente a las estrategias irrevocables de búsqueda binaria, cual de las siguientes es CORRECTA**

- No permite la vuelta atrás.

**31. ¿Que estructura tiene un sistema de producción?**

- Base de conocimiento y Motor de inferencia.

**32. ¿Cuál de estas opciones es correcta?:**

- La desventaja principal de la heurística es que requiere un mayor coste computacional.

**33. Utilizando la técnica de relajación de la restricción de optimalidad “Algoritmo de ponderación dinámica”. ¿Qué tipo de búsqueda se realiza en los últimos niveles?:**

- Ninguna es correcta

**34. Un algoritmo A1 es dominante sobre A2, si...**

- cada nodo expandido por A1 es también expandido por A2.

**35. De las siguientes opciones, ¿cuál es la incorrecta sobre el mantenimiento de la admisibilidad?:**

- Consumir poco espacio.

**36. ¿Cuál es el objetivo de la técnica de admisibilidad- $\epsilon$ ?**

- Aumentar la velocidad de búsqueda a costa de obtener una solución subóptima.

**37. Estrategias de búsqueda básica que para aplicarlas sea necesario que no importe el camino al objetivo:**

- Ambas son correctas.

**38. ¿Cuál es el principal problema de la búsqueda de estados?**

- Introducir una componente inteligente en el proceso, la heurística.

**39. ¿Qué problemas podemos resolver mediante técnicas de búsqueda?**

- Ambas respuestas son correctas.

**40. De la terna (BH, RP, EC) de un sistema de producción, indica la definición correcta:**

- RP (Reglas de Producción): Conjunto de operadores para la transformación de los estados del problema, es decir, de la base de hechos. Cada regla tiene dos partes: Precondiciones; Postcondiciones;

**41. ¿Es posible aplicar la distancia de Manhattan en un 8-con?**

- No, porque la diagonal ya va a ser mínima.

**42. Acerca de heurística se puede decir que:**

- En general el nivel de información de las heurísticas permite encontrar antes la solución, pero tiene la desventaja de requerir un mayor coste computacional para su cálculo.

**44. A\* $\epsilon$  opera de forma idéntica al algoritmo A\* salvo que selecciona aquel nodo de Lista\_Focal...**

- con menor valor de  $H_f(n)$ .

**45. Sabemos que la estrategia de control (EC) determina el conjunto de reglas aplicables mediante un proceso de pattern-matching y resuelve conflictos entre varias reglas a aplicar mediante el filtrado. Dependiendo del caso ¿cómo podemos escoger la EC más adecuada?**

- Identificando si se puede descomponer el problema ante el que nos encontramos, viendo si se pueden ignorar o al menos deshacer pasos erróneos hacia la solución, descubriendo si el universo es predecible, comprobando si la bondad de una solución es relativa o absoluta y si la solución es un estado o un camino. Además, ver el papel que desempeña el conocimiento a la hora de encontrar la solución.

**46. En la resolución de teoremas, ¿qué tipo de pasos hacia la solución son los adecuados durante la elección de la Estrategia de Control si se confirman que son erróneos?**

- Ignorables.

**47. Clasificando las estrategias de búsqueda básicas, podemos afirmar que:**

- Se dividen en Irrevocables y Tentativas.

**48. Los conceptos básicos de la búsqueda heurística son:**

- Compleitud, admisibilidad, dominación y optimalidad.

**49. En cuanto a las estrategias de búsqueda basica tentativas, sabemos que...**

- Se mantienen estados de vuelta atrás por si el estado actual no llega a buen fin.

**50. Según la formalización de los problemas de búsqueda, realizado por POST en 1943, un sistema de producción Base de Hechos (BH) es:**

- Conjunto de representaciones de uno o más estados por los que atraviesa el problema. Constituye la estructura de datos global.

**51. Si comparamos tres algoritmos para resolver un problema, A, B y C, evaluando estos el 100%, el 80% y el 60% de la totalidad de nodos generables para ese problema, respectivamente. Podemos afirmar que:**

- C es el algoritmo óptimo.

**52. Para un problema de camino más corto en el que nos podemos mover en las 8 direcciones, podemos afirmar que la heurística más óptima es:**

- Cálculo de la diagonal.

**53. ¿Cuál de estas estrategias de búsqueda no permite la vuelta atrás?:**

- Irrevocable.

**54. En grafos infinitos con meta alcanzable, ¿podremos obtener una solución óptima?**

- Sí, tanto con búsqueda en anchura como con coste uniforme, siendo más óptima esta última.

**55. El coste de una función de búsqueda de un camino mínimo y óptimo es  $f(n) = g(n) + h(n)$  de forma que:**

- si tenemos una heurística  $h(n)$  cuyo valor es constante o nula para cada nodo "n" del recorrido, cualquier nodo "n" del mismo nivel tendrá el mismo coste.

**56. ¿Dado un problema de camino mínimo en el que sólo podemos movernos en 4 direcciones (arriba, abajo, izquierda, derecha), cual de las siguientes es la función de la heurística óptima (siendo x e y nuestra posición y m y n las coordenadas objetivo)?**

- $h^*((x,y)) = |m-x| + |n-y|$

**57. ¿Cuál de las siguientes afirmaciones es verdadera?**

- El uso de una heurística  $h(n) = 0$  admite la posibilidad de encontrar una solución óptima.

**58. ¿Cómo se define la heurística en nuestra asignatura?**

- La función que estima lo que nos va a costar llegar a la solución óptima.

**59. ¿Cuando se dice que una función heurística  $h(n)$  es admisible?**

- Cuando la función heurística garantiza la obtención de un camino de coste mínimo hasta un objetivo.

**60. En una Búsqueda irrevocable se requiere:**

- Disponer del suficiente conocimiento local.

**61. Ayuda a la familia a cruzar el puente. Ten en cuenta que es de noche y necesitan la linterna para cruzar. - Cada miembro cruza a una velocidad distinta:**

- A: 1s
- B: 3s
- C: 6s
- D: 8s
- E: 12s

**- El puente sólo resiste un máximo de 2 personas. - Un par debe cruzar a la velocidad del miembro más lento. - La linterna sólo dura 30s.**

**Aplicando las siguientes reglas de producción:**

- R1 - pareja cruza el puente
- R2 - vuelve el individuo.

**indica la solución al problema:**

- $(AB - R1) (A - R2) (DE - R1) (B - R2) (AC - R1) (A - R2) (AB - R1)$ .

**62. La BH (Base de Hechos) podría definirse como:**

- Conjunto de representaciones de uno o más estados por los que atraviesa el problema.

**63. ¿Cual de las siguientes afirmaciones es correcta?**

- Ninguna de las anteriores es correcta.

**64. La búsqueda en profundidad...**

- Ninguna de las anteriores es correcta.

**65. ¿Qué problemas tienen las estrategias de búsqueda irrevocables?**

- Tiene problemas de mesetas, máximos locales y crestas.

**66. El algoritmo de búsqueda A\* utiliza una función de evaluación  $f(n) = g(n) + h'(n)$ , en la que  $h'(n)$  representa el valor heurístico del nodo a evaluar, desde el nodo actual,n, hasta el nodo terminal.**

**¿Cuándo esta función ( $h'(n)$ ) no es admisible?**

- Si se sobreestiman los gastos reales para alcanzar el nodo terminal.

**67. ¿Cuál de las siguientes afirmaciones respecto a la clasificación de estrategias de búsqueda básica es correcta?**

- Las dos anteriores son correctas.

**68. En el Ciclo de Control Básico en una estrategia de control...**

- Se define por cuatro estados. E1:Exploración de la frontera, E2:Cálculo de reglas aplicables, E3:Resolución de conflictos y E4:Aplicación.

**69. ¿Puede la búsqueda de costo uniforme volverse infinita?**

- Si cuando el coste del nodo expandido sea cero y conduzca de nuevo al mismo estado.

**70. Si he diseñado una estrategia heurística para encontrar el camino a la meta en un laberinto, solo con esos datos, puedo afirmar que .. :**

- Es una estrategia tentativa.

**71. Teniendo en cuenta las siguientes definiciones:**

[ $G^*(n)$ ]: Coste del camino de coste mínimo desde el nodo inicial s al nodo n. [ $C^*$ ]: coste del camino mínimo desde el nodo inicial a un nodo solución. [S]: Estado inicial de partida. [Tj]: Estado final o meta. [n]: nodo intermedio. [ $H^*(n)$ ]: Coste del camino de coste mínimo de todos los caminos desde el nodo n a cualquier estado solución Tj. [ $F^*(n)$ ]: Coste del camino de coste mínimo desde el nodo inicial hasta un nodo solución condicionado a pasar por n.

**Indica cual de las siguientes afirmaciones es verdadera:**

- 1. $F^*(n)=C^*$  en cada nodo del camino óptimo.

**72. A la hora de hallar la heurística para un problema, debemos tener en cuenta:**

- Las dos son correctas.

**73. Un granjero se encuentra en la orilla de un río junto con un lobo, una cabra y una col. Dispone de un bote en el que sólo puede transportar una única cosa cada vez. El granjero pretende transportar al lobo, la cabra y la col al otro lado del río utilizando el bote. No puede dejar solos en una orilla al lobo y a la cabra, ni a la cabra y la col. ¿Cómo conseguiría el granjero trasladar todo a la margen derecha del río?**

Suponiendo el espacio de estados: (Pg, Pl, Po, Pc)

- Pg = 1 o 2
- Pl = 1 o 2
- Po = 1 o 2
- Pc = 1 o 2
- Pg representa la orilla del río donde está el granjero.
- Pl representa la orilla del río donde está el lobo.
- Po representa la orilla del río donde está la oveja.
- Pc representa la orilla del río donde está la col.
- 1 representa orilla inicial.
- 2 representa orilla final

Dadas las siguientes reglas, elige la secuencia que resuelve el problema cumpliendo con todas las restricciones planteadas en el enunciado:

Regla	(Pg, Pl, Po, Pc)	(Pg, Pl, Po, Pc)
1	Si (1, 1, 1, 1)	→ (2, 1, 2, 1)
2	Si (1, 1, 1, 1)	→ (2, 2, 1, 1)
3	Si (1, 1, 1, 1)	→ (2, 1, 1, 2)
4	Si (2, 1, 1, 1)	→ (1, 2, 1, 1)
5	Si (2, 1, 1, 2)	→ (1, 1, 1, 2)
6	Si (2, 1, 2, 1)	→ (1, 1, 2, 1)
7	Si (1, 1, 1, 2)	→ (2, 1, 2, 2)
8	Si (1, 1, 2, 1)	→ (2, 2, 2, 1)
9	Si (2, 1, 1, 1)	→ (2, 2, 2, 2)
10	Si (2, 1, 1, 2)	→ (1, 1, 1, 2)
11	Si (2, 1, 2, 2)	→ (1, 1, 2, 1)
12	Si (2, 2, 2, 1)	→ (1, 2, 1, 1)
13	Si (1, 2, 1, 1)	→ (2, 2, 1, 2)
14	Si (1, 1, 2, 1)	→ (2, 1, 2, 2)
15	Si (2, 1, 2, 2)	→ (1, 1, 2, 2)
16	Si (2, 2, 1, 2)	→ (1, 2, 1, 2)
17	Si (2, 2, 2, 1)	→ (1, 1, 1, 2)
18	Si (1, 2, 2, 1)	→ (2, 2, 2, 2)
19	Si (1, 2, 1, 2)	→ (2, 2, 2, 2)
20	Si (1, 1, 2, 2)	→ (2, 2, 2, 2)

- Reglas: 1, 6, 8, 12, 13, 16, 19

**74. El algoritmo de ponderación dinámica es aplicable donde:**

- Se conoce la Profundidad.

**75. Un algoritmo A\* debe su nombre a:**

- Una función heurística admisible.

**76. Los requerimientos que deben cumplir las estrategias irrevocables y tentativas son:**

- Todas son correctas.

**77. ¿Cual de las siguientes afirmaciones sobre las estrategias tentativas desinformadas es correcta?**

- Búsqueda en anchura y coste uniforme serán similares cuando el coste de aplicación de cada regla sea unitario.

**78. ¿Que tres partes tiene el sistema de producción de POST?**

- BH, RP, EC.

**79. En una búsqueda irrevocable por gradiente de un mínimo, puede ser un problema encontrar:**

- Una meseta.

**80. En la búsqueda heurística podemos afirmar que:**

- Nos interesa encontrar algoritmos de tipo A admisibles.

**81. La estrategia de búsqueda ciega es mas eficiente en...**

- pequeños problemas.

**82. En un algoritmo A (de aditivos),  $g(n)$  en un nodo no final es la estimación:**

- del coste del camino de coste mínimo desde el estado inicial hasta el nodo n.

**83. En cuanto a la estrategia tentativa no informadas, selecciona la opción correcta:**

- Son ciegas en el sentido de que el orden en el cual la búsqueda progresá no depende de la naturaleza de la solución que buscamos.

**84. Dentro de las estrategias tentativas no informadas, ¿en cuál no se garantiza obtener una solución óptima en grafos finitos?**

- En la búsqueda en profundidad.

**85. Teniendo en cuenta las estrategias de búsqueda básicas, cual de estas afirmaciones es cierta:**

- En las estrategias Irrevocables el avance debe ser metódico , mantenemos una frontera unitaria y tenemos presente que nunca ahí vuelta atrás.

**86. En cuanto a la subdivisión las estrategias de búsqueda básicas, las estrategias:**

- Tentativas: La búsqueda es multi o mono camino. Se mantiene estados de vuelta atrás por si el estado actual no llega a buen fin.

**87. En la ponderación de programación dinámica....solo es aplicable en:**

- Ambas.

**88. En un algoritmo B que utiliza una función heurística admisible...:**

- si nos devuelve un valor superior a  $h^*$ , para algún nodo, no se puede garantizar que encontremos la solución óptima.

**89. Conociendo:**

$$h1((x, y)) = |m-x| + |n-y| \quad h2((x,y)) = |x| + |y| \quad h3((x,y)) = \sqrt{(m-x)^2 + (n-y)^2}$$

siendo m y n las coordenadas x e y del destino.

**En un problema de camino mínimo en el que nos movemos en 4 direcciones:**

- $h_1$  domina sobre  $h_2$ .

**90. ¿En qué caso es admisible un algoritmo de ponderación dinámica?**

- En problemas donde se conoce la profundidad en la cual va a aparecer la solución.

**91. Siendo una función heurística  $h(n)$ :**

- Decimos que es admisible si obtiene óptimos y caminos de coste mínimo.

**92. Si consideramos el problema del viajante de comercio, una heurística admisible sería:**

- Calcular la distancia utilizando el algoritmo de Dijkstra

**93. Cual NO es un inconveniente de mantener la admisibilidad:**

- No es práctico en problemas pequeños

**94. ¿Cual de las siguientes afirmaciones es cierta?**

- Todo algoritmo óptimo es completo.

**95. Respecto a la técnica de ponderación dinámica o APD: ¿Cuál es el motivo por el que la búsqueda se frena en niveles cercanos a la solución?**

- Porque el entorno de admisibilidad tiende a cero.

**96. A la hora de elegir una heurística, debemos tener cuidado de no...:**

- dejar fuera al nodo óptimo.

**97. En el ejemplo de heurísticas para el 8-puzzle tenemos que la mejor aproximación heurística a la hora de encontrar solución al problema es:**

- La suma de las distancias de las piezas a sus posiciones en el objetivo utilizando la distancia Manhattan.

**98. Los inconvenientes de mantener la admisibilidad en las estrategias heurísticas son aquellos que fuerzan al algoritmo a consumir mucho tiempo en discriminar caminos cuyos costes no varían muy significativamente y que se consume mucha memoria ya que se mantienen los nodos generados en memoria. ¿Cuál de las siguientes opciones no representa una posible solución al problema?**

- Aumentar la velocidad para obtener una solución óptima.

**99. ¿Cuál es el objetivo de la técnica de admisibilidad- $\epsilon$ ?**

- Aumentar la velocidad de búsqueda a costa de obtener una solución subóptima.

**100. ¿Cuál de estas estrategias de búsqueda no permite la vuelta atrás?:**

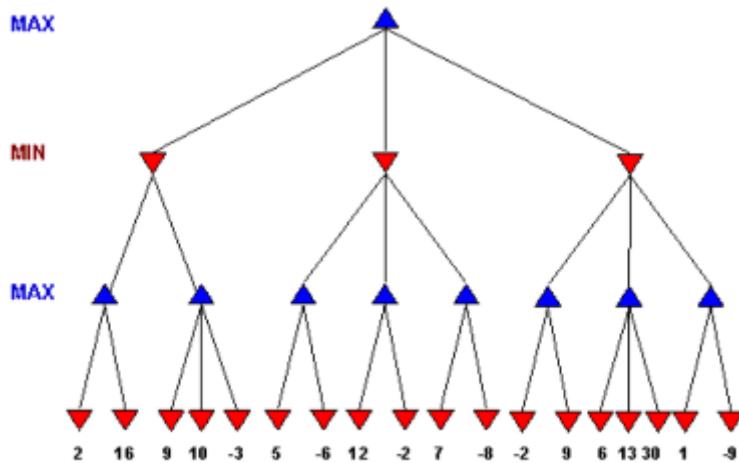
- Irrevocable.

### Preguntas T3: Búsqueda en juegos

1. En juegos con restricciones temporales de respuesta, la técnica de bajada progresiva, consiste en:

= Recorrer los nodos por niveles y al llegar la petición (final del tiempo) devolver la solución del último nivel completado.

2. Dado el siguiente árbol de juego:



¿Cuál será el valor MINIMAX del nodo raíz?

= 10

3. Indica cuál de las siguientes afirmaciones es correcta:

= La poda alfa beta es una técnica de búsqueda que reduce el número de nodos evaluados en un árbol de juego por el Minimax.

4. La estrategia Exhaustiva MiniMax

= La raíz siempre debe tener valor

5. La espera en reposo busca evitar:

~ Expandir todos los nodos.

~ Llegar a una situación estable en la que encontrar la mejor solución

= Ninguna de las anteriores

## **6. En la poda heurística**

~ La función adicional de evaluación  $g(N)$  tiene un alto coste

~ Se intenta evitar el efecto horizonte

= Se reduce  $B$ (factor de ramificación) desarrollando únicamente los mejores movimientos de cada nivel

## **7. ¿Para los juegos multijugador?:**

=se sustituye el valor de un nodo por un vector de valores (tantos como jugadores)

## **8. La función $g(N)$ de la poda heurística tiene que cumplir la condición de:**

Tener un coste computacional pequeño.

## **9. Una de las características principales de la estrategia MiniMax es:**

Generar todos los nodos hasta la profundidad deseada

## **10. El efecto horizonte es:**

Dado un problema sólo tenemos conocimiento hasta una profundidad dada.

## **11. El rendimiento de un programa se puede mejorar si... :**

Se le proporcionan movimientos de libro

## **12. Respecto a la Continuación Heurística:**

Selecciona un subconjunto de nodos terminales para desarrollar búsquedas más profundas.

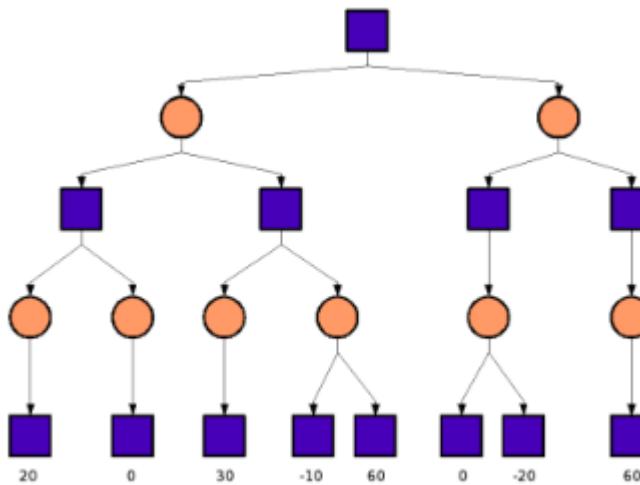
## **13. En la estrategia de poda: $a-\beta$ : {**

~  $a$  es el valor de la mejor opción hasta el momento a lo largo del camino para MIN, esto implicará por lo tanto la elección del valor más bajo.

~ $\beta$  es el valor de la mejor opción hasta el momento a lo largo del camino para MAX, esto implicará por lo tanto la elección del valor más alto.

= Las dos anteriores son falsas.

**14. En el siguiente arbol minimax de un juego de 2 personas, en el que los cuadrados son nodos max y los circulos nodos min.**



**El resultado es:**

- 20

**15. ¿Cuándo se nos podría presentar un problema aplicando la estrategia MiniMax?:**

= Cuando necesitamos una respuesta en muy poco tiempo.

**16. Estamos utilizando una estrategia de poda  $\alpha\text{-}\beta$  y nos encontramos en un nodo MAX, el valor de  $\alpha$  (mejor opción de max) es 9, ¿con qué valor podaríamos las ramas restantes?:**

= 7

**17. En la técnica de bajada progresiva, el factor de ramificación se calcula como:**

El factor de ramificación no pertenece a la técnica complementaria nombrada.

**18. El factor de ramificación es:**

El número de posibles movimientos que se pueden realizar.

**19. El uso de movimientos de libro:**

= Se usa durante el inicio y final de la partida, y se combina con MiniMax para la parte central de la partida.

**20. ¿Cuál de las siguientes afirmaciones NO es correcta acerca de la estrategia exhaustiva Minimax:**

= Permite evitar las restricciones de tiempo, pudiendo devolver un resultado correcto en el momento de petición de jugada.

**21. En el supuesto de una partida de ajedrez que se encuentre en la situación inicial, sin que se haya hecho ningún movimiento, ¿Qué técnica complementaria convendría utilizar, a parte de la de minimax, para que la computadora ejecutase el primer movimiento?**

= La técnica de movimiento de libro.

**22. En la técnica de bajada progresiva:**

= Se recorren los nodos por niveles.

**23. En la estrategia de poda  $\alpha$ - $\beta$  los valores iniciales son:**

=  $\alpha = -\infty$ ;  $\beta = \infty$

**24. ¿En estrategia de poda  $\alpha$ -  $\beta$  cuando se realizara la poda de las ramas restantes?**

=cuando el valor actual que se está examinando sea peor que el valor actual de  $\alpha$  para MAX o  $\beta$  para MIN.

**25. Con la estrategia exhaustiva MiniMax, se quiere conseguir:**

= determinar valor del nodo raíz.

**26. ¿Porque se ha de usar los llamados movimientos de libro?:**

=Porque es imposible seleccionar un movimiento consultando la configuración actual del juego en un catálogo y extrayendo el movimiento correcto.

**27. En cuanto a la búsqueda avanzada y juegos, ¿qué técnicas serían complementarias?{**

~ Uso de movimientos de libro.

~ Poda heurística.

= Ambas son correctas.

**28. La técnica de espera del reposo consiste en..**

= explorar nodos hasta que su valor no cambie de manera drástica después de explorar un nivel más, es decir, el valor del nodo se estabilice de un nivel al siguiente.

**29. En la estrategia de poda:  $\alpha$ - $\beta$  debemos empezar ... :**

= Inicializando  $\alpha = -\infty$ ;  $\beta = \infty$ .

**30. En búsqueda en juegos, cuál de las técnicas complementarias hace referencia a la aplicación de una serie movimientos predefinidos dependiendo de la situación en determinado instante del juego, con el fin de mejorar el rendimiento del algoritmo:**

= Uso de movimientos de libro

**31. ¿Cuál de las siguientes técnicas de búsqueda utiliza una estrategia exhaustiva? {**

= MiniMax.

**32. En la estrategia de poda  $\alpha$ - $\beta$ ....**

= las 2 son correctas

**33. ¿El valor MiniMax de un nodo estará siempre acotado por...?: {**

=  $\alpha \leq V(N) \leq \beta$

**34. En los juegos como problemas de búsqueda:**

= En cada nivel se van alternando los jugadores.

**35. En la poda heurística, ¿cuál es el factor de ramificación?:**

~ Factor(Nodo)= Rango(Nodo) - Factor(Padre(Nodo))

~ Factor(Nodo)= Factor(Padre(Nodo)) + Rango(Nodo)

= Ninguna de las anteriores

**36. La espera del reposo busca: {**

= Evitar el efecto horizonte.

**37. ¿Qué es el estado de un árbol?{**

= Configuración del juego en un momento dado

**38. En las estrategias de búsqueda avanzada y juegos, cuando hablamos de la estrategia exhaustiva MiniMax, indica la respuesta correcta:**

= Al valor que determina al nodo raíz, se le denomina valor MiniMax.

**39. ¿Por qué la búsqueda en juegos requiere un tratamiento especial?:{**

= Porque es imposible generar todo el árbol de búsqueda.

**40. ¿Cuándo es aconsejable el uso de movimientos de libro?**

= Para ciertas partes de juegos que están muy estudiadas

**41. En la Estrategia exhaustiva MiniMax es FALSO que:**

= Minimax es extendido a juegos de solo un jugador.

**42. El objetivo de la poda heurística es:{**

= reducir B (factor de ramificación) desarrollando únicamente los mejores movimientos de cada nivel.

**43. En los juegos Multijugador podemos afirmar que no es cierto que:**

= Usamos Minimax extendido a juegos de dos jugadores como mucho

**44. En la poda heurística, consideramos que la función g(N):{**

= Una versión simplificada de f(N).

**45. La estrategia de poda alfa beta es:**

= es una técnica de búsqueda que reduce el número de nodos evaluados en un árbol de juego por el Minimax.

**46.** Si al expandir el nodo correspondiente al estado actual del juego obtenemos 4 nodos hijos. ¿Cuántos nodos se expandiran en el siguiente nivel si estamos usando poda heuristica?{

=6

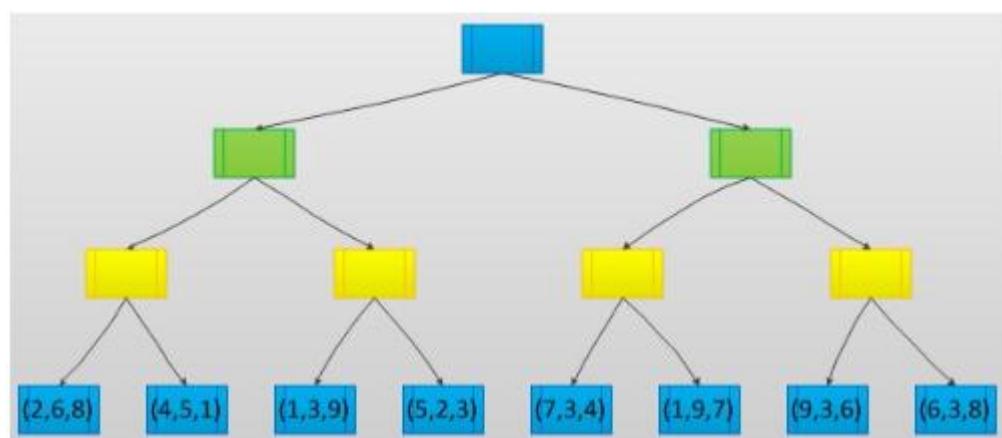
**47.** El uso de movimientos de libro nos permite:

= Alcanzar una solución analizando el estado actual del juego y aplicando unos movimientos concretos almacenados en un catálogo.

**48.** La técnica de bajada progresiva permite:

= Recorrer nodos por nivel y al llegar a la petición de jugada, devolver la solución del último nivel que se halla completado.

**49.** Dado el siguiente árbol de multijugador (Azul- jugador 1; Verde – jugador 2; Amarillo – jugador 3):



¿Cúal será el valor MinMax del nodo raiz?

= (2,6,8)

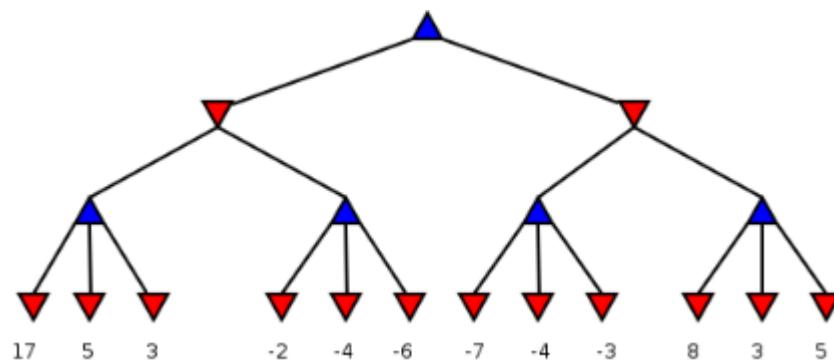
**50.** El uso de movimientos de libro en una partida de ajedrez:

= Se usan en las aperturas y los finales de las partidas, combinado con el procedimiento MiniMax para la parte central de la partida

**51. De los siguientes pasos, ¿cuál de ellos no pertenece a la estrategia exhaustiva MiniMax?**

= Seleccionar un subconjunto de nodos terminales para desarrollar búsquedas más profundas

**52. Dado el siguiente árbol de juego y teniendo en cuenta que los coloreados en rojo corresponden a los nodos MIN y los azules a los nodos MAX:**



**¿Qué valor tomará el nodo raíz?**

= -2

**53. Una búsqueda exhaustiva hasta una profundidad deseada:**

= consiste en una búsqueda por niveles hasta llegar al nivel deseado.

**54. El uso de movimiento de libro:**

= Se usa el libro en las aperturas y en los finales combinado con el MiniMax

**55. Estamos empleando poda heurística para resolver un juego. En cierto punto existen n posibles movimientos. ¿Cuál será el factor del primer nodo de profundidad n que nos encontraremos?**

= 0

**56. ¿Qué es la poda alfa beta?**

= Técnica de búsqueda que reduce el número de nodos evaluados en un árbol de juego por el Minimax.

**57. Efecto horizonte es la limitación en profundidad, solo se puede tener conocimiento hasta la profundidad seleccionada.**

= Correcto, todo el árbol tiene un determinado nivel de profundidad.

**58. En el análisis de un arbol de juegos de busqueda las aristas representan:**

= Un posible movimiento.

**59. ¿Por qué en el juego del ajedrez no se puede aplicar la búsqueda exhaustiva?**

= Porque es imposible generar todo el árbol de búsqueda.

**60. Wolfgang Kempelen crea el "ajedrecista mecánico" en el año: {**

= 1760

**61. Escoge la opción correcta respecto a la estrategia MiniMax:**

= Una mayor profundidad en el arbol implica una mayor probabilidad de ganar, ya que reduce el factor horizonte

**62. La poda α-β...:**

= reduce el número de nodos evaluados en un árbol de juego por el Minimax.

**63. ¿Cuál de las siguientes afirmaciones sobre técnicas complementarias de búsquedas en juegos es correcta?**

= El objetivo de la poda heurística es reducir B desarrollando únicamente los mejores movimientos de cada nivel.

**64. Dada la técnica de bajada progresiva es cierto que:**

= Tiene una restricción de tiempo para resolver el árbol devolviendo la mejor solución que le ha dado tiempo a explorar

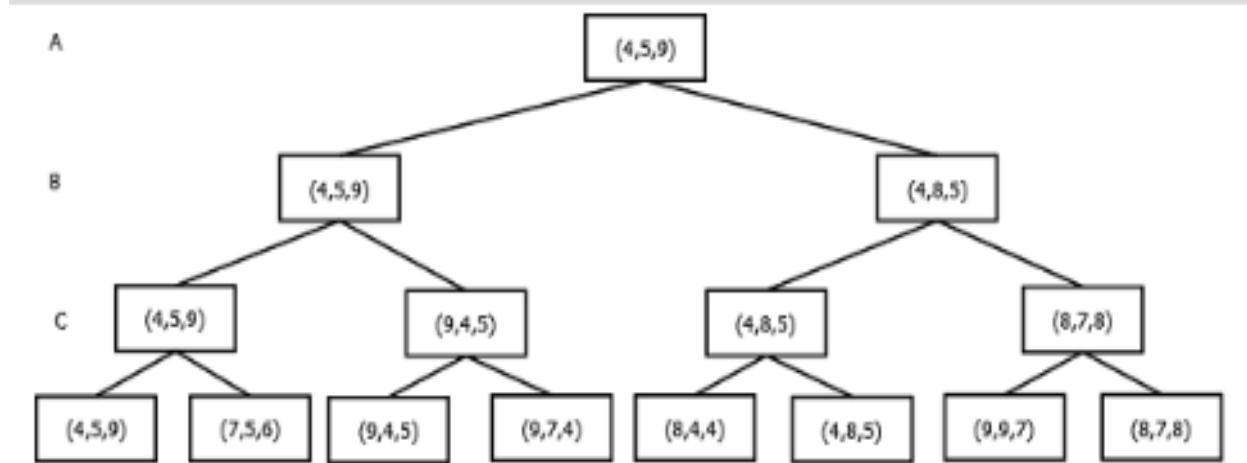
**65. ¿Cómo se puede moderar el “efecto horizonte” en la teoría de los juegos?**

= Con la espera del reposo

**66. La técnica de bajada progresiva consiste:**

= El recorrido de los nodos es por niveles y se devuelve la solución del ultimo nivel que se haya completado.

**67. Dado el siguiente árbol de juegos con tres jugadores (usando MAX):**



**Indica cuál de las siguientes afirmaciones es correcta.**

= A se alía con C (coalición).

**68. Selecciona la respuesta correcta:**

= En la poda Heurística reducir B desarrollando únicamente los mejores

movimientos de cada nivel, siendo el factor de ramificación:  
Factor(Nodo) = Factor(Padre(Nodo)) - Rango(Nodo)

**69. En juegos como problemas de búsqueda, qué opción es incorrecta:**

= Arbol de juego: cada arista de ese árbol indica un posible movimiento. Una rama completa contempla todas las posibles jugadas

**70. El objetivo de la poda heurística es:**

= reducir el factor de ramificación B desarrollando únicamente los mejores movimientos de cada nivel

**71. Selecciona la respuesta correcta:**

= Uso de movimientos de libro: *El rendimiento del programa puede mejorarse si se le proporciona una lista de movimientos.*

**72. Escoge la opción correcta respecto a las técnicas complementarias de búsquedas en juegos:**

=En ajedrez, tanto la secuencia de apertura como los finales están muy estudiados, por lo tanto podríamos usar el uso de movimientos de libro en estas partes del juego.

**73. En un juego con restricción de tiempo a la hora de decidir la jugada conviene:**

= Aplicar una técnica de bajada progresiva.

**74. ¿De que manera mejora la técnica α-β la estrategia Minimax?**

= la técnica α-β reduce el número de nodos evaluados de forma que puede distinguir que partes del árbol no van a mejorar el resultado

**75. Teniendo en cuenta que ciertos juegos podemos abordarlos como problemas de búsqueda:**

= El inicio de la jugada queda definido por el análisis del árbol.

**76. La poda alfa beta es realizada...**

=Durante la construcción del árbol

**77. ¿Puede usarse Minimax en juegos multijugador?**

= Si, puede utilizarse para cualquier número de jugadores.

**78. ¿Que es el efecto horizonte?**

= Es cuando al no poder evaluar más allá de cierto punto elige una mala opción

**79. El efecto horizonte (continuacion heuristica) esta provocado por:**

= La limitación en profundidad: solo se puede tener conocimiento hasta la profundidad seleccionada.

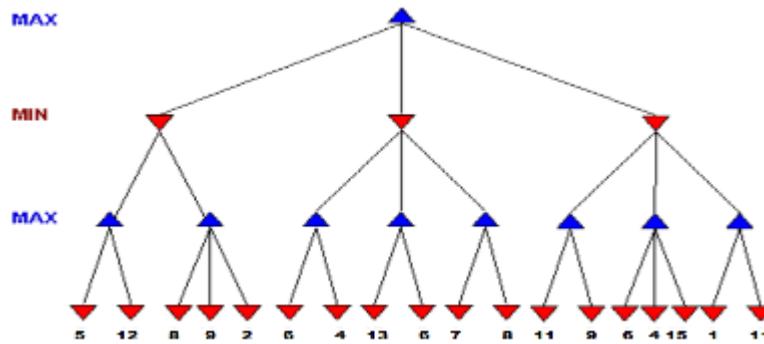
**80. ¿Qué es el factor de ramificación?**

= Número de posibles movimientos que se pueden realizar.

**81. Cual de las siguientes afirmaciones es verdadera:**

= La estrategia de poda  $\alpha\beta$  es la única que hace reducciones de nodos utilizando MinMax, inicializamos  $\alpha = -\infty$  cuando no tenemos ninguna evidencia y el recorrido inicia en la raíz y terminamos en las hojas.

**82. ¿Cuáles serán los valores MINIMAX del nodo raíz, así como  $\alpha$  y  $\beta$ ?**



= MINIMAX = 11,  $\alpha = 15$ ,  $\beta = 1$

**83. De las siguientes técnicas para mejorar la estrategia exhaustiva MiniMax, cuál es la que está enfocada en superar las limitaciones del tiempo:**

= Técnica de bajada progresiva.

**84. En la búsqueda en juegos, varias de las técnicas complementarias estudiadas son:**

= Espera del reposo, poda heurística, y continuación heurística.

**85. Cuáles de las siguientes características hace referencia a la técnica complementaria "Técnica de bajada progresiva":**

= Se utiliza en juegos donde existan restricciones de tiempo. Al llegar a la petición de jugada se devuelve la solución del último nivel completado.

**86. En juegos con poda  $g(N)$  tiene:**

= Ninguna de las anteriores es correcta.

**87. ¿Qué llamamos "movimientos de libro" y qué efecto tienen en un programa?**

= Son movimientos estudiados que deben realizarse en ciertos casos, y sirven para mejorar el rendimiento del programa.

**88. ¿Cuándo se ejecutará la poda en la búsqueda alfa beta de las ramas restantes?**

= Cuando el valor que se está examinando sea peor que el valor actual de  $\alpha$  o  $\beta$  para MAX o MIN

**89. Se plantea el diseño de un algoritmo para jugar al ajedrez y se nos requiere implementar un MiniMax. De las siguientes opciones, ¿Qué estrategia sería mejor para obtener un programa potente sin gran coste computacional?**

= Utilizando un libro de movimientos que nos aporte una gran base de conocimiento.

**90. Decimos que la técnica de bajada progresiva:**

= Obtiene el nodo solución del último nivel completamente recorrido del árbol de soluciones

**91. La técnica de bajada progresiva:**

= Recorre el árbol por niveles y cuando se acaba el tiempo devuelve la solución del último nivel explorado.

**92. La efectividad del método de poda alfa-beta depende en mayor medida:**

= Del orden en el que los nodos sucesores son examinados.

**93. En la búsqueda en juegos, centrándonos en la técnica de espera del reposo, podemos afirmar que...:**

= Si un nodo cambia su valor de manera drástica después de haber explorado un nuevo nivel, la búsqueda debe continuar.

**94. La técnica complementaria a la estrategia de poda  $\alpha - \beta$  denominada "Técnica de bajada progresiva" es el más adecuada cuando disponemos de:**

= Restricciones de tiempo

#### T4ejemplo 1

Dadas la variables  $x, y, z$  con dominios

$$Dx = Dy = \{1, 2, 3, 4, 5\}, Dz = \{0, 1\}$$

Con restricciones  $x \leq y-1, y \geq z+4, x = z + 3$ . ¿Qué respuesta es cierta?: {

~ Aplicando el algoritmo AC3, los dominios restringidos que cumplen las consistencias de arco finales son:

$$CDx = \{3, 4\}$$

$$CDz = \{1\}$$

$$CDy = \{4, 5\}$$

= Aplicando el algoritmo AC3, los dominios restringidos que cumplen las consistencias de arco finales son:

$$CDx = \{3, 4\}$$

$$CDz = \{0, 1\}$$

$$CDy = \{4, 5\}$$

~ Aplicando el algoritmo AC3, los dominios restringidos que cumplen las consistencias de arco finales son:

$$CDx = \{3, 4\}$$

$$CDz = \{0, 1\}$$

$$CDy = \{3, 4\}$$

}

Explicación:

Los dominios que cumplen las restricciones binarias son:

$$\text{para } x \leq y-1 \quad x \in \{1, 2, 3, 4\} \quad y \in \{2, 3, 4, 5\}$$

$$\text{para } y \geq z+4 \quad y \in \{4, 5\} \quad z \in \{0, 1\}$$

$$\text{para } x = z + 3 \quad x \in \{3, 4\} \quad z \in \{0, 1\}$$

Aplicando AC3: (eliminamos inconsistencias) obtenemos los dominios restringidos:

$$CDx = \{3, 4\}$$

$$CDz = \{0, 1\}$$

$$CDy = \{4, 5\}$$

Siendo las dos soluciones posibles:

$$z=0 \quad x=3 \quad y=4$$

$$z=1 \quad x=4 \quad y=5$$

#### T4pregunta495\_1

Tratándose de Criptoaritmética, una de sus características es: {

- ~ Dominios discretos y restricciones simples.
- = Dominios discretos y restricciones múltiples.
- ~ Dominios complejos y restricciones simples.

}

Explicación: Como podemos comprobar en el tema (transparencia 10\_T4) se refiere a una característica de la Criptoaritmética (Sustituir cada letra por un dígito distinto de manera que la suma sea correcta).

#### T4pregunta879\_1

Según las limitaciones del Backtracking, la inconsistencia de arista: {

- ~ Está relacionado con las restricciones unarias.
- = Está relacionado con las restricciones binarias.
- ~ Ambas son correctas.

}

Explicación: En la trasparencia 18 del T4. Sucede cuando existe una restricción binaria entre dos variables de tal forma que para un determinado valor de la primera variable no existe ninguna asignación posible para la segunda.

#### T4pregunta948\_1

¿En qué consiste la búsqueda mediante backtracking?{

- ~ Se trata de construir la solución de forma no gradual.
- = Se trata de construir la solución de forma gradual, instanciando variables en el orden definido por la permutación dada
- ~ Ninguna de las anteriores

}

Explicación: Como podemos ver en la transparencia 13 del tema 4 la respuesta correcta es la segunda.

#### T4pregunta659\_1

Un ejemplo de CSP binario es: {

- ~ Generación de crucigramas.
- ~ Coloreado de mapas.
- = Los dos ejemplos anteriores son correctos.

}

Ejemplos de CSP binarios: coloreado de mapas, asignación de tareas para un robot, N-reinas, generación de crucigramas, etc.

#### T4pregunta091\_1

Según Forward checking{

= Los valores de las variables futuras que son inconsistentes con la asignación actual son temporalmente eliminados de sus dominios

~ Los valores de las variables futuras que son inconsistentes con la asignación actual son eliminados de sus dominios

~ Ninguna de las anteriores

}

Explicación: Como bien viene explicado en el tema, son eliminadas temporalmente porque es posible que luego sean utilizados para la solución

#### T4pregunta417\_1

¿Cuáles son las características del problema de las N-Reinas? :{

- ~ Dominios continuos y restricciones binarias.
- = Dominios discretos y restricciones binarias.
- ~ Dominios discretos y restricciones múltiples.

}

Explicación: Las respuestas erróneas corresponden respectivamente a la Criptoaritmética y a las Restricciones temporales.

#### T4pregunta136\_1

En un tablero de ajedrez de tamaño  $n \times n$  y con  $n >= 4$ , el máximo número de reinas que se pueden colocar sin que se ataquen entre ellas es: {

- ~  $n - 1$
- =  $n$
- ~  $n + 1$

}

Explicación: en un tablero de ajedrez de tamaño  $n \times n$  con  $n >= 4$  se puede colocar hasta una reina por fila consiguiendo que no se ataquen entre ellas.

#### T4pregunta614\_1

Sobre el CSP binario se puede afirmar que:{

- ~ El problema de las n-reinas no se puede formular como un CSP binario.
- = Todo problema n-ario se puede formular como un problema binario.
- ~ Las dos anteriores son falsas.}

Explicación: Según la descripción del CSP binario todo problema n-ario se puede formular como un problema binario. Es por esto que el problema de las n-reinas sí que se puede formular como un CSP Binario.

#### T4pregunta564\_1

¿Cuál de las siguientes afirmaciones es o son limitaciones del método Backtracking?: {

~ Trashing y dependencia de valor.

= Inconsistencia de arista.

~ Inconsistencia de variable.

}

Explicación: Inconsistencia de arista. Ya que puede ocurrir cuando existe una restricción binaria entre 2 variables y no hay forma de asignar un valor a la segunda para un determinado valor de la primera porque no existe dicho valor posible. De ahí la inconsistencia.

#### T4pregunta283\_1

¿Qué nos permite la propagación de restricciones?: {

~ Transformar el problema a uno más sencillo, pero manteniendo las inconsistencias de arco.

= Transformar el problema a uno más sencillo sin inconsistencias de arco.

~ Ninguna de las anteriores.

}

Explicación: Tal y como podemos apreciar en la transparencia número 22 del tema 4, propagamos las restricciones con la finalidad de evitar una inconsistencia que nos dificulte la resolución del problema.

#### T4pregunta237\_1

¿Cómo se construye la solución en un árbol de interpretaciones?: {

- = De forma incremental en la que cada hoja es una interpretación
- ~ De forma decremental en la que cada hoja es una interpretación
- ~ Ninguna de las anteriores.

}

*Explicación: En un árbol de interpretaciones partimos de un nodo raíz en el que se supervisa el proceso, siendo cada nivel una asignación para una característica de datos, es decir, por cada nivel es una asignación parcial, siendo en el último nivel donde se realiza la asignación total.*

#### T4pregunta948\_1

Para la resolución de problemas de gran tamaño, ¿en qué orden de mayor eficiencia a menor eficiencia, se situarían los siguientes métodos de resolución? {

- ~ Generación y Test > Backtracking > Backjumping
  - ~ Backtracking > Backjumping > Generación y Test
  - = Backjumping > Backtracking > Generación y Test
- }

*Explicación: El método menos eficiente de los tres es el de Generación y Test, ya que no aplica ninguna técnica de poda. Entre Backtracking y Backjumping, la diferencia es que Backjumping cuando encuentra un espacio de dominios vacíos salta hasta el nodo de conflicto, no hasta la ramificación anterior como hace Backtracking, por lo tanto con Backjumping se pueden evitar los errores que puede provocar Backtracking al ir expandiendo ramas anteriores.*

#### T4Pregunta275\_1

¿Mediante que algoritmo se puede convertir un CSP en una red consistente?: {

~ AC4

= AC3

~ AC2.

}

Explicación: Tal y como podemos apreciar en la transparencia número 22 del tema 4, mediante este sencillo algoritmo se puede realizar esta función.

#### T4Pregunta667\_1

En que se basa el algoritmo AC3:

{

~Es un algoritmo que elimina las restricciones para buscar una solución

~Examina las aristas, creando una red consistente en el dominio de cada variable

=Examina las aristas, eliminando los valores que causan inconsistencia del domino de cada variable

}

Explicación: El algoritmo no crea uniones sino que las elimina las que son inconsistentes.

#### T4pregunta347\_1

En el método Forward checking si el dominio de una variable futura se queda vacío:

{

= La instanciación de la variable actual se deshace y se prueba con un nuevo valor.

~ Comprueba hacia delante la asignación actual con todos los valores.

~ Ninguna de las anteriores.

}

Explicación: Si el dominio de una variable futura se queda vacío, la instanciación de la variable actual se deshace y se prueba con un nuevo valor. Si ningún valor es consistente, entonces se lleva a cabo el backtracking cronológico.

#### T4pregunta000\_1

Una arista dirigida  $c(ep) = \langle V_i, V_j \rangle$  es consistente si y sólo si: {

- = para todo valor asignable a  $V_i$  existe al menos un valor en  $V_j$  que satisface la restricción asociada a la arista.
  - ~ para ningún valor asignable a  $V_i$  existe al menos un valor en  $V_j$  que satisface la restricción asociada a la arista.
  - ~ para todo valor asignable a  $V_i$  existe al menos un valor en  $V_i$  que satisface la restricción asociada a la arista.
- }

Explicación: Según la definición que podemos encontrar en la página 22 de los apuntes del tema: Una arista dirigida  $c(ep) = \langle V_i, V_j \rangle$  es consistente si y sólo si para todo valor asignable a  $V_i$  existe al menos un valor en  $V_j$  que satisface la restricción asociada a la arista.

#### T4pregunta176\_1

Como se puede representar un CSP: {

- = Ninguna de las anteriores.
  - ~ Un vector.
  - ~ Una matriz.
- }

Explicación: Un CSP se puede representar como un grafo, por eso la respuesta correcta es la primera.

#### T4pregunta715\_1

Acerca de las soluciones de CSP puede decir que:{

- ~ Un buen sistema CSP normalmente encontrar una solución.
  - ~ No todos los problemas N-arias se pueden convertir en problemas binarios, para la simplificación de la misma.
  - = Ninguna de las anteriores es correcta.
- }

*Explicación: La primera alternativa es falsa porque es posible encontrar más de una solución, en función de las necesidades, o incluso sin solución.*

*La segunda también es falsa, los problemas complejos se puede simplificar aún si implica en una alteración del modelo inicial.*

#### T4pregunta327\_1

El CSP es:{

- ~ Conjunto de variables definidas sobre dominios infinitos y conjunto de restricciones definidas sobre subconjuntos de dichas variables.
  - = Conjunto de variables definidas sobre dominios finitos y conjunto de restricciones definidas sobre subconjuntos de dichas variables.
  - ~ Conjunto de variables definidas sobre dominios finitos y sin restricciones definidas sobre subconjuntos de dichas variables.
- }

*Explicación: El conjunto de variables tiene que tener un dominio finito y además tiene que tener restricciones sobre los subconjuntos de esas variables.*

#### T4pregunta760\_1

En un árbol de interpretaciones:{

- = Partimos de un nodo raíz que supervisa el proceso.
- ~ La solución se construye de forma decremental de tal forma que cada hoja es una interpretación.
- ~ Todas son correctas.

}

*Explicación: La primera opción es la correcta porque la segunda para también serlo debería de ser, La solución se construye de forma incremental, no decremental, de tal forma que cada hoja es una interpretación. Por esto no todas son correctas.*

#### T4pregunta470\_1

Si utilizamos Propagación de restricciones, obtendremos: {

- ~ Al menos una solución
- = Una, ninguna o varias soluciones
- ~ Una o ninguna solución

}

Si hacemos uso de la propagación de restricciones, al finalizar, obtendremos ninguna (no tiene solución), varias (puesto que las tuplas de los nodos se han quedado con más de un elemento) o solo una, que sería la solución directa.

#### T4pregunta344\_1

construir la solución de forma gradual, instanciando variables en el orden definido por la permutación dada, es un método de resolución de:

- {
- ~ Generación y test
- = Backtracking
- ~ Ninguna de las anteriores es correcta.
- }

Explicación: La diferencia entre backtracking y generación y test es que generación y test buscan por todas las soluciones del problema. Mientras que lo backtracking se consigue construyendo soluciones parciales a medida que progresó el recorrido; estas soluciones parciales limitan las regiones en las que se puede encontrar una solución completa.

#### T4Pregunta512\_1

Cuál de las siguientes afirmaciones NO es correcta acerca de la estrategia de generación y test:{

- = Transforma el problema en uno más sencillo, el cual, es testeado para comprobar su consistencia.
- ~ Es muy poco eficiente
- ~ Se basa en expandir una a una todas las posibilidades del problema.
- }

Explicación: La estrategia de generación y test se basa en comprobar el espacio de soluciones al completo, y de ello deriva su baja eficiencia. La transformación del problema en más sencillos pertenece a la estrategia de propagación de restricciones, y no se testeaa nada sino que se desarrolla la estrategia a base de comprobación de consistencias.

#### T4pregunta718\_1

¿Cuál es la única diferencia que encontramos entre Backtracking y Backjumping?{

- ~ Backtracking siempre encuentra la solución óptima y Backjumping no.
- = Cuando encontramos un espacio de dominios vacío, Backtracking sólo puede volver al nodo anterior, es decir, subir un nivel mientras que Backjumping puede saltar al nodo en conflicto.
- ~ Ninguna de las anteriores.

}

Explicación: Backtracking cuando decide volver hacia atrás, sólo puede retroceder al nodo padre y de ahí seguir la búsqueda de la solución óptima, en cambio, Backjumping tiene la capacidad de saltar directamente al nodo en conflicto y ahorrar algunos pasos, esto es más eficiente ante un problema grande.

#### T4Pregunta907\_1

¿Cuáles de estos métodos de resolución es híbrido?{

- ~ Backtracking.
- = Forward checking.
- ~ Backjumping.

}

Explicación: Categorías en las transparencias 13 y 14 del tema 4, en el caso de que en las comprobaciones previas no tiene valores consistentes, FC realiza un backtracking

#### T4\_Pregunta363\_1

En el problema del coloreado de mapas:{

- ~ Se tiene que resolver siempre utilizando el algoritmo AC3, ya que backtracking no nos garantiza que vaya a encontrar una solución.
- ~ Se necesita el mismo número de colores que número de territorios fronterizos con el área que más territorios fronterizos tenga.
- = Ambas son incorrectas

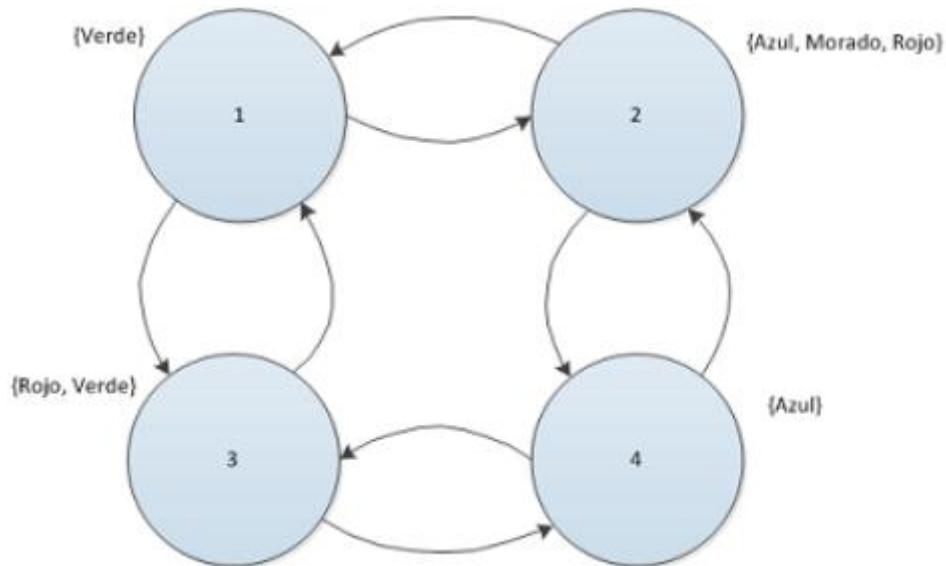
}

La primera es incorrecta porque este problema se puede resolver por backtracking, otra cosa es que tarde una cantidad de tiempo inasumible.

La segunda es incorrecta porque se necesita un color más, que es el que le das al propio territorio, o si no el problema no tendrá solución.

#### T4pregunta911\_1

Aplicando la regla  $p_k (V_i, V_j) = \{<V_i, V_j> | V_i \in D_i, V_j \in D_j, V_i \neq V_j\}, \forall k, 1 \leq k \leq 4$ , ¿Qué aristas dirigidas del siguiente grafo no serían consistentes en el instante inicial?



```
{  
~ (1, 3), (2, 1)  
= (3, 1), (2, 4)  
~ (4, 3), (4, 2)  
}
```

Explicación: Según la propiedad de consistencia de arista: "Una arista dirigida  $c(ep) = <V_i, V_j>$  es consistente si y sólo si para todo valor asignable a  $V_i$  existe al menos un valor en  $V_j$  que satisface la restricción asociada a la arista."

Podemos observar que en el caso de la arista  $(3, 1)$ , para  $V_i = \text{Rojo}$  tendríamos  $\langle \text{rojo}, \text{verde} \rangle$ , pero en el caso de  $V_i = \text{Verde}$  la única opción posible sería  $\langle \text{verde}, \text{verde} \rangle$  que no cumpliría  $v_i \neq v_j$ , lo mismo ocurre para el caso de la arista  $(2, 4)$  en el caso de  $V_i = \text{Azul}$

Las otras opciones son correctas ya que si nos fijamos por ejemplo en el caso  $(1, 3)$  tendríamos  $\langle \text{verde}, \text{rojo} \rangle$  o  $\langle \text{verde}, \text{verde} \rangle$  pero la propiedad indica que existe **al menos un valor en  $V_j$  que satisfaga la restricción**, en este caso sería el rojo.

#### T4pregunta532\_1

¿Cuáles son las características de las restricciones temporales?{

- ~ Dominios discretos y restricciones binarias.
- ~ Dominios discretos y restricciones múltiples.
- = Dominios continuos y restricciones binarias.

}

#### T4pregunta660\_1

Las heurísticas se clasifican como métodos de resolución por:{

- ~ Inferencia
- ~ Búsqueda
- = Algoritmos híbridos

}

Explicación: Las heurísticas se clasifican como métodos de resolución mediante algoritmos híbridos, ya que sobre un esquema de búsqueda incorporan métodos de inferencia.

#### T4pregunta406\_1

Dadas la variables  $x, y, z$  con dominios

$$Dx = \{0, 1, 2\}$$

$$Dy = \{3, 4\}$$

$$Dz = \{0, 1\}$$

Con restricciones:

$$x \geq y - 1$$

$$y \geq z + 4$$

$$x = z + 2$$

¿Cuál de las siguientes afirmaciones es cierta?: {

~ Aplicando el algoritmo AC3, podemos asegurar que NO tiene solución 

= Aplicando el algoritmo AC3, podemos asegurar que tiene una única solución

~ Aplicando el algoritmo AC3, podemos asegurar que tiene más de una solución

}

Explicación:

Los dominios que cumplen las restricciones binarias son:

$$\text{para } x \geq y - 1 \quad x \in \{2\} \quad y \in \{3, 4\}$$

$$\text{para } y \geq z + 4 \quad y \in \{4\} \quad z \in \{0\}$$

$$\text{para } x = z + 2 \quad x \in \{2\} \quad z \in \{0\}$$

Aplicamos AC3 eliminando inconsistencias, obteniendo los siguientes dominios restringidos:

$$CDx = \{2\}$$

$$CDz = \{0\}$$

$$CDy = \{4\}$$

Por lo que la única solución posible es:

$$x = 2$$

$$y = 4$$

$$z = 0$$

#### T4pregunta053\_1

Sobre las limitaciones de Backtracking, cuando hay una restricción binaria entre dos variables hablamos de... {  
= Inconsistencia de arista  
~ Inconsistencia de nodo  
~ Ninguna de las anteriores  
}

Explicación: La correcta es "inconsistencia de arista" ya que un determinado valor de una variable no tiene una asignación posible con la segunda variable. Por ejemplo.

Tenemos dos nodos adyacentes (Grafo dirigido) y entre  $\langle x,y \rangle$  y  $\langle y,x \rangle$  cumplen una misma restricción R1:  $x \leq y * 4$

Los dominios de X e Y son los siguientes:

Dx: {1,2,3}

Dy: {1,2}

Si lo comprobamos, veremos que la restricción no se cumple en ningún caso.

#### T4\_pregunta475\_1

En Backtracking, si No se puede extender la solución parcial:{

- ~ Se elimina la última decisión.
- ~ Se elimina una decisión anterior.
- = Las dos anteriores son correctas.

}

Explicación: según sea cronológico o no cronológico la 2 primeras opciones son correctas. diapositiva 16 tema 4

#### T4Pregunta306\_1

¿Cuál de las siguientes afirmaciones de Generación y test es correcta?{

- = Busca la solución mediante una expansión del árbol en anchura.
- ~ Busca la solución mediante una expansión del árbol en altura.
- ~ Ignora ciertas soluciones para alcanzar un óptimo.

}

Explicación: Generación y test comprueba el espacio de soluciones en su totalidad, por eso en un nodo genera todas las posibilidades posibles.

#### T4pregunta044\_1

Tiene sentido que Backtracking calcule la expansión de todos los nodos:{

- ~ Sí, para tener todas las soluciones.
- ~ Sí, porque el orden de aplicación de las acciones es conmutativo.
- = No, porque hay conmutatividad}

Explicación: Las asignaciones de variables son conmutativas, por ejemplo para dos zonas en un mapa, Alicante y Valencia, las asignaciones Alicante = rojo, Valencia = amarillo equivalen a Valencia = amarillo, Alicante = rojo, con lo que calcular dos veces (en este ejemplo) lo mismo no tiene sentido.

#### T4pregunta668\_1

En el algoritmo AC3, cuando un dominio queda vacío ¿qué significa?. Que el problema es: {

- = Inconsistente y sin solución.
- ~ Consistente y sin solución.
- ~ Consistente y con solución única.

}

Explicación: cuando un dominio queda vacío no existen valores de las variables que cumplan las restricciones.

#### T4pregunta212\_1

He cometido una equivocación, la pregunta correcta es:

Las características de la Generación de crucigramas son :{

- ~ Dominios continuos y restricciones binarias.
- ~ Dominios discretos y restricciones múltiples.
- = Dominios discretos y CSP binario.

}

Explicación: La generación de crucigramas no tiene conjunto continuo de dominios, es discreto (dominios grandes), y tiene restricciones cada dos slots (CSP binario).

Las respuestas que no son válidas corresponden a las características de las restricciones temporales y a las de la criptoaritmética.

#### T4pregunta480\_1

Dentro de los métodos de solución, en el backtracking no cronológico {

- ~ Se elimina la última decisión.
- = Se elimina la decisión anterior.
- ~ No elimina ninguna decisión.

}

*Explicación: Si en el método de backtracking no se puede extender la solución parcial que se construye hasta una solución total, pasamos al backtracking cronológico en el que se elimina la última decisión, o bien al backtracking no cronológico donde se descarta la decisión anterior.*

#### T4pregunta418\_1

El backjumping {

- ~ hace el retroceso a la primera variable instanciada
- ~ no hace retroceso
- = hace el retroceso a la variable más profunda que está en conflicto con la variable actual.

}

*Explicación: La opción a) sólo sería válida si la primera variable instanciada fuera la que está en conflicto con la actual, pero esto no podemos saberlo con anterioridad.*

#### T4pregunta 736\_1

¿Cuáles son las características de la generación de crucigramas?

{

- ~Dominios continuos y restricciones ternarias.
- ~CSP ternario y sin restricciones.
- =CSP Binario, discreto (dominios grandes)

}

*Explicación: Ver tema 4 página 8.*

#### T4pregunta850\_1

De entre las siguientes limitaciones del método Backtracking selecciona aquella que es incorrecta:{

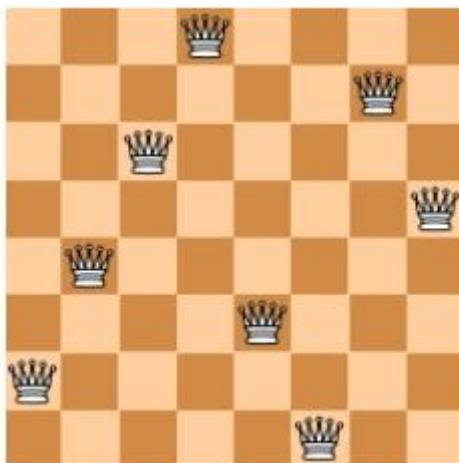
- ~ Trashing e inconsistencia del nodo. Sucede cuando un dominio contiene un valor que no satisface una restricción unaria.
- = Consistencia de arista. Sucede cuando una restricción binaria entre dos variables de tal forma que para un determinado valor de la primera variable existe una asignación posible para la segunda.
- ~ Dependencia de ordenación. El orden de selección de las variables es un factor crítico. Existen diferentes heurísticas de selección de variable y de valor, esto es, variables con Orden Estático y Orden Dinámico y valores p.e que conducen a CPS más simples.

}

Explicación: la consistencia de arista precisamente no es una limitación para Backtraking ya que las relaciones binarias serían correctas y parejas, en cambio cuando para un determinado valor de una variable no existe ninguna asignación posible de otra variable se crea una inconsistencia de arista que sí que es una limitación.

#### T4pregunta606\_1

Dada la siguiente solución en el problema de las 8reinas:



Partiendo de que la reina  $X_i$  está la fila  $i$ -ésima, ¿qué representación es correcta? {

~  $(7, 5, 3, 1, 6, 8, 2, 4)$

~  $(4, 7, 8, 3, 2, 1, 5, 6)$

=  $(4, 7, 3, 8, 2, 5, 1, 6)$

}

Explicación: La primera no es correcta porque corresponde con la representación en la columna  $i$ -ésima. La segunda no es correcta porque en la fila 3 no hay reina en la columna 8 y así con el resto de filas. La tercera es correcta porque representa tal cual las posiciones actuales de las reinas.

#### T4pregunta536\_1

El Forward checking, en cada etapa de la búsqueda, comprueba hacia delante la asignación actual con {

~un valor de las futuras variables que están restringidas con la variable actual.

~uno o varios valores de las futuras variables que están restringidas con la variable actual.

=todos los valores de las futuras variables que están restringidas con la variable actual.

}

Explicación: según la transparencia 19 del tema 4; forward checking es un algoritmo que se basa en comprobar todos los valores de las futuras variables.

#### T4pregunta093\_1

¿Qué diferencia hay entre Backjumping y Backtracking? {

~ Backjumping retrocede hacia la variable más profunda que está en conflicto con la actual, Backtracking realiza el retroceso hacia la variable más prometedora.

= Backjumping retrocede hacia la variable más profunda que está en conflicto con la actual, Backtracking realiza el retroceso hacia la variable instanciada anteriormente.

~ Ninguna. Ambas nomenclaturas sirven para definir el mismo método de resolución.

}

Explicación: Backtracking y Backjumping nombran dos métodos de resolución distintos, el retroceso de Backtracking no es retroceder hacia la variable más prometedora, Backtracking retrocede hacia la variable anteriormente instanciada y Backjumping retrocede hasta la variable más profunda en conflicto con la actual.

#### T4pregunta290\_1

Centrándonos en la propagación de restricciones y propiedad de consistencia de arista podemos afirmar que es correcto que: {

~ Un CSP no puede transformarse en una red consistente mediante un algoritmo sencillo (AC3).

~ La propagación de restricciones no se suele usar porque transforma el problema en otro más complejo con inconsistencias de arco.

= Un CSP puede transformarse en una red consistente mediante un algoritmo sencillo (AC3).

}

Explicación: según la transparencia 22 del tema 4; Un CSP puede transformarse en una red consistente mediante un algoritmo sencillo (AC3) que examina las aristas, eliminando los valores que causan inconsistencia del dominio de cada variable..

#### T4pregunta150\_1

Cuál de las siguientes afirmaciones es verdadera:

~ en el Árbol de interpretaciones cada nodo identifica una posibilidad de asignación, cada hoja es una interpretación de la solución

y esta misma supervisa todo el proceso comparando su resultado volviendo a el nodo raíz .

=en el árbol de interpretaciones el orden de descenso viene especificado por a, todos los nodo se evalúa si puede existir

una posibilidad de asignación según su valor y va haber tantos niveles como asignaciones de datos.

~ el nodo raíz supervisa el proceso de asignación, solo se especifica una posibilidad de especificación cuando hemos expandido todo

el árbol en profundidad asignando una variable por cada nivel del árbol.

Explicación:

1º no es correcta por que la hoja no es la que supervisa todo el proceso si no el nodo a.

3º no es correcta porque una posibilidad de asignación se hace en cada nivel por esto mismo es incremental.

#### T4pregunta798\_1

Cuál de las siguientes opciones es correcta: {

~ Backjumping construye la solución de forma gradual y su retroceso se hace a la variable instanciada anteriormente.

~ Backtracking construye la solución de forma gradual y su retroceso se hace a la variable más profunda que está en conflicto con la variable actual.

= Generación y test genera de forma sistemática y exhaustiva cada una de las posibles asignaciones y comprueba que satisfacen todas las restricciones.

}

Explicación: Las definiciones de backtraking y backjumping están intercambiadas, por lo que la correcta es la última. Se puede ver en la diapositiva 13 del tema 4

#### T4pregunta187\_1

Indica cuál de las siguientes afirmaciones es cierta sobre al algoritmo AC3:{

- ~ Si el grafo es consistente obligatoriamente tiene que tener solución
  - = Si el grafo es consistente puede tener una solución o más
  - ~ Si el grafo es inconsistente podemos encontrar solución
- }

Explicación: Un grafo consistente no indica obligatoriamente que tenga solución, puede tenerla o no, por ello puede tener una solución pero no obligatoriamente tenerla.

#### T4pregunta322\_1

En lo referente a los métodos de resolución, podemos afirmar que Generación y test: {

- ~ Construye la solución de forma gradual, instanciando variables en el orden definido por la permutación dada.
  - ~ Construye la solución de forma gradual, instanciando la variable más profunda que está en conflicto con la variable actual.
  - = Genera cada una de las posibles asignaciones a las variables y comprueba si satisfacen todas las restricciones.
- }

Explicación: Mediante Generación y test generamos cada una de las posibles asignaciones a las variables y comprobamos si satisfacen todas las restricciones, además, exploramos el espacio definido por el producto cartesiano de los dominios de las variables.

#### T4pregunta730\_1

- Si hablamos de Forward checking podemos afirmar que: {
- ~ Los valores de las variables futuras que son inconsistentes con la asignación actual no son eliminados de sus dominios.
  - ~ Si ningún valor es consistente, entonces se lleva a cabo el backtracking no cronológico
  - = Ninguna de las anteriores
- }

Explicación: Los valores futuros que son inconsistentes con la asignación actual en forward check se descartan y si no se encuentra ningún valor consistente se aplica el backtraking cronológico.

#### T4pregunta609\_1

Dentro de los métodos de resolución tenemos los Algoritmos híbridos. ¿Cuál de los siguientes es un algoritmo híbrido?: {

- ~ Forward Checking.
  - ~ Maintaining Arc Consistency.
  - = Las respuestas anteriores son ambas correctas.
- }

Explicación: En la transparencia 14 del Tema 4 podemos comprobar que tanto el algoritmo "Forward Checking" como el "Maintaining Arc Consistency" pertenecen a este grupo.

#### T4pregunta492\_1

CSP busca: {

- ~ La relación N-aria que satisface todas las restricciones del problema.
  - ~ Segundo los requerimientos del problema hay que encontrar todas las soluciones o sólo una.
  - = Ambas son correctas.
- }

Explicación: Ambos son metas del CSP cogidas de las transparencias.

#### T4pregunta363\_1

¿Qué método de búsqueda para problemas de satisfacción de restricciones, construye la solución de forma gradual, instanciando variables en el orden definido por la permutación dada? {

- ~ Generación y test
  - ~ Backjumping
  - = Backtracking
- }

*Explicación: Por definición la respuesta correcta es Backtracking*

#### T4pregunta072\_1

Tras la utilización del algoritmo AC3 hemos logrado eliminar todos los valores que causan inconsistencia del dominio de cada variable. Con este exitoso resultado podemos afirmar que:{

- ~ Como hemos logrado eliminar toda inconsistencia no nos hará falta hacer backtracking nunca.
  - ~ Tendremos que concluir utilizando backtracking siempre.
  - = Usaremos backtracking únicamente cuando la eliminación de inconsistencias de lugar a más de una solución
- }

*Explicación: Basarse en la consistencia de arco no es suficiente para garantizar evitar el uso de backtracking*

#### T4pregunta917\_1

En búsqueda para problemas de satisfacción de restricciones las limitaciones de "inconsistencia de arista", "trashing e inconsistencia de nodo" y "dependencia de la ordenación" pertenecen al método de resolución:{

- ~ Esquema de propagación de restricciones.
- = Esquema backtracking.
- ~ Esquema forward checking.

}

Explicación: Las Limitaciones del backtracking son:

Inconsistencia de arista: Relacionado con las restricciones binarias. Sucede cuando existe una restricción binaria entre dos variables de tal forma que para un determinado valor de la primera variable no existe ninguna asignación posible para la segunda.

Trashing e inconsistencia de nodo: Relacionado con las restricciones unarias.

Sucede cuando un dominio contiene un valor que no satisface una restricción unaria.

Dependencia de la ordenación: El orden de selección de las variables es un factor crítico. Se han desarrollado diversas heurísticas de selección de variable y de valor.

#### T4pregunta366\_1

¿Cuál de entre los siguientes podría considerarse como método de inferencia puro?{

- ~ Heurísticas.
- = Consistencia de caminos.
- ~ Forward Checking.

}

Explicación: Puede observarse en la diapositiva 14 del tema 4. Tanto las heurísticas como el forward checking son algoritmos híbridos.

#### T4pregunta591\_1

En un problema de satisfacción de restricciones (CSP){

~ Cada restricción implica al conjunto de variables y especifica las combinaciones aceptables de valores para dicho conjunto.

~ La solución es una relación n-aria que satisface algunas restricciones del problema.

=Visualizarlo como un grafo de restricciones puede usarse para simplificar el proceso de solución.

Explicación: La primera opción es falsa: Ver definición de un CSP (Un conjunto de restricciones definidas sobre **subconjuntos** de dichas variables). La segunda opción es falsa: La solución siempre debe satisfacer **todas las restricciones** del problema.

La tercera opción es verdadera, al representar un CSP como un grafo de restricciones simplifica el proceso de solución, produciendo en algunos casos una reducción exponencial de la complejidad (ver Inteligencia Artificial Un enfoque moderno 5. Problemas de satisfacción de restricciones)

#### T4pregunta516\_1

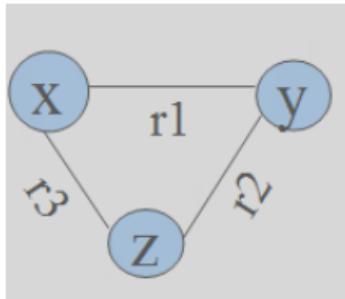
Siendo:

$$D(x) = \{1,2,3,4,5,6\} \quad D(y) = \{1,2,3,4,5\} \quad D(z) = \{2,3,4,6\}$$

$$r1: x \geq y + 1$$

$$r2: y > z + 1$$

$$r3: z \leq x - 2$$



Aplicando el algoritmo AC3, indica la respuesta correcta:

- {
- =  $x = 5, y = 4, z = 3$ . Es una posible solución
- $\sim x = 3, y = 1, z = 4$ . Es una posible solución
- $\sim x = 5, y = 4, z = 3$ . Es solución única.

}

Explicación:

Aplicamos r1  $x > y + 1 \rightarrow D(x)=\{2,3,4,5,6\} \quad D(y)=\{2,3,4,5\}$

Aplicamos r2  $y > z + 1 \rightarrow D(y)=\{4,5\} \quad D(z)=\{3,4,5\}$

Aplicamos r3  $z \leq x - 2 \rightarrow D(z)=\{3,4\} \quad D(x)=\{5,6\}$

Obtenemos:

$D(x) = \{5,6\} \quad D(y) = \{4,5\} \quad D(z) = \{3,4\}$

por lo que una posible solución sería:

$x = 5 \quad y = 4 \quad z = 3$

T4pregunta270\_1

Un CSP binario es aquel: {

- $\sim$  En el que toda las variables son discretas, es decir, toman valores en dominio discretos.

- = En el que todas las restricciones tienen a los sumo dos variables respectivamente.

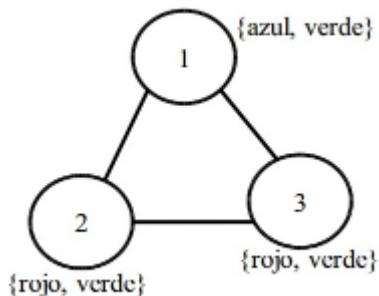
- $\sim$  En que constan de variables continuas y discretas.

}

La definición de un CPS Binario es "Un CSP binario es aquel en el que todas las restricciones tienen a los sumo dos variables respectivamente".

#### T4pregunta411\_1

Podemos afirmar que el grafo de la imagen es un grafo: {



- ~ Es consistente sin solución.
- = Consistente con dos soluciones.
- ~ Inconsistente con dos soluciones.

}

Explicación: Que un grafo tenga solución garantiza que sea consistente, pero que sea consistente no garantiza que tenga solución. En este caso es un ejemplo claro de consistente con dos soluciones.

#### T4pregunta217\_1

Respecto al Forward checking... {

- ~ En cada etapa de la búsqueda, comprueba hacia atrás la asignación actual con algunos de los valores de las futuras variables que están restringidas con la variable actual.

- ~ En cada etapa de la búsqueda, comprueba hacia adelante la asignación actual. con algunos de los valores de las futuras variables que están restringidas con la variable actual.

- = Ninguna de las anteriores es correcta.

}

Explicación: En cada etapa de la búsqueda, Forward checking comprueba hacia delante la asignación actual con todos los valores de las futuras variables que están restringidas con la variable actual.

#### T4pregunta028\_1

- Un CSP puede transformarse en una red consistente mediante u  
algoritmo sencillo (AC3)... {  
= examinando las aristas, eliminando los valores que causan la inconsistencia del  
dominio de cada variable.  
~ eliminando las aristas, examinando los valores de cada variable.  
~ ninguna de las anteriores.  
}

Explicación: Tema 4, trasparencia 22.

#### T4pregunta336\_1

- Cuál de los pares de métodos de resolución corresponde a un método de inferencia  
seguido de un método híbrido:{  
~ Maintaining Arc Consistency, Consistencia de arcos.  
~ Forward Checking, K-consistencia.  
= Consistencia de caminos, Heurísticas.  
}

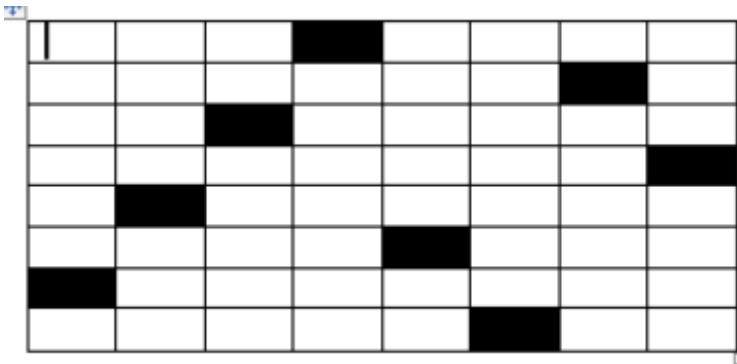
Explicación: Los métodos de resolución por inferencia son Consistencia de arcos,  
Consistencia de caminos y K-consistencia. Los métodos híbridos son Maintaining Arc  
Consistency, Forward Checking y Heurísticas.

#### T4pregunta210\_1

- ¿Cuál sería una solución para el problema de n-reinas en un tablero de 4x4? {  
~ (1, 3, 4, 2)  
= (3, 1, 4, 2)  
~ (4, 1, 3, 2)  
}

Explicación: Si para cada opción colocamos cada una de las reinas en la casilla  
correspondiente nos damos cuenta de que solo la segunda opción cumple los  
requisitos de que ninguna reina se encuentre en la misma columna o en diagonal a  
otra reina.

#### T4pregunta567\_1



Dado un tablero de 8x8 y el problema de las reinas, ¿Cuál es el vector que satisface la solución del tablero mostrado? :{

$$= S = (6, 4, 2, 0, 5, 7, 1, 3)$$

$$\sim S = (6, 2, 4, 0, 5, 7, 1, 3)$$

$$\sim S = (3, 1, 4, 0, 5, 7, 1, 3)$$

}

Explicación = La posición del vector indica la columna en la que está la reina y el numero la fila.

#### T4pregunta897\_1

Hallar el máximo valor que puede tomar la palabra "HOLA" si :

$$\begin{array}{r} \text{HHH} \\ + \text{O} \\ \hline \end{array} \quad H \neq O \neq L \neq A \neq 0$$

$$\sim 8579$$

$$\sim 9859$$

$$= 8759$$

}

Explicación: El ejercicio de criptoaritmética se resuelve de la siguiente manera, como HOLA debe tomar el máximo valor, o sea H debe tomar su máximo valor H = 9, pero si hacemos la operación respectiva notamos que no cumple (999 + x = yyyy), ahora hacemos H = 8, O=5, L=7, A=9 y vemos que 8+5 != 7, por lo tanto HOLA = 8759 porque 888 + 7 = 895

#### T4pregunta218\_1

Usando el algoritmo AC3 transforme en una red consistente:

Variables:  $V = \{X, Y, Z\}$

Dominios:  $Dx = Dy = \{1, 2, 3, 4, 5\}$

$Dz = \{0, 1, 2\}$

Restricciones:  $\rho_1 \quad Y \leq X - 1$

$\rho_2 \quad X \geq Z + 4$

$\rho_3 \quad Y = Z + 2$

Después de aplicar el algoritmo, como son los dominios de cada variable?{

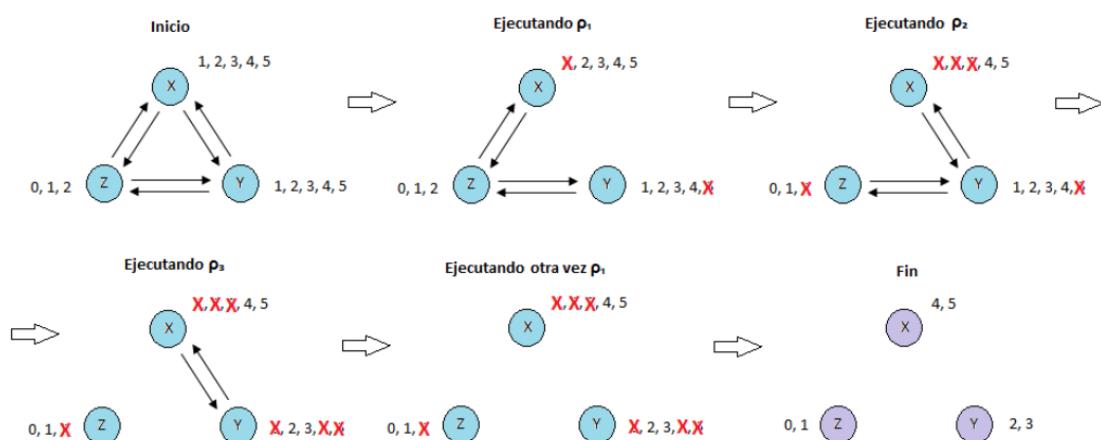
=  $Dx = \{4, 5\}; Dy = \{2, 3\}; Dz = \{0, 1\}$

~  $Dx = \{1, 2, 4, 5\}; Dy = \{1, 2, 3\}; Dz = \{0, 1\}$

~  $Dx = \{1, 2, 3\}; Dy = \{4, 5\}; Dz = \{1, 2\}$

}

*Explicación:* Transformarse la red consistente mediante el algoritmo sencillo (AC3) que examina las aristas, eliminando los valores que causan inconsistencia del dominio de cada variable. Aplicación del algoritmo:



#### T4pregunta279\_1

La clase más simple de problemas de satisfacción de restricciones implica variables discretas y dominios finitos. Sin embargo las variables discretas pueden tener también dominios infinitos (conjunto de números enteros o de cadenas, etc).

Respecto a estos últimos indica que afirmación es correcta:

{

- ~ Con dominios infinitos es posible describir restricciones enumerando todas las combinaciones permitidas de valores.
- ~ Con dominios infinitos es necesario usar lenguajes de restricción y resolver las restricciones enumerando las asignaciones posibles.

= Ninguna de las anteriores es correcta.

}

Explicación:

La primera afirmación es falsa ya que hablamos de combinaciones de valores infinitas, es imposible enumerarlas todas. La segunda también lo es por la misma razón, a pesar de usar lenguajes de restricción no podemos enumerar todas las asignaciones posibles porque siguen siendo infinitas, debemos acotar en los casos donde se pueda los valores que toman las variables. Por tanto ninguna de las afirmaciones es correcta.

#### T4pregunta405\_1

Dadas la variables  $x, y, z$  con dominios

$$Dx = \{1, 3, 5\}, Dy = \{2, 4, 6\}, Dz = \{1, 2, 3\}$$

Con restricciones  $x \geq y+3, y \geq z, z \leq x-3$ . ¿Qué respuesta es cierta?: {

= Aplicando el algoritmo AC3, los dominios restringidos que cumplen las consistencias de arco finales son:

$$CDx = \{5\}$$

$$CDz = \{2\}$$

$$CDy = \{1, 2\}$$

~ Aplicando el algoritmo AC3, los dominios restringidos que cumplen las consistencias de arco finales son:

$$CDx = \{5\}$$

$$CDz = \{2\}$$

$$CDy = \{2\}$$

~ Aplicando el algoritmo AC3, los dominios restringidos que cumplen las consistencias de arco finales son:

$$CDx = \{5\}$$

$$CDz = \{2\}$$

$$CDy = \{1\}$$

}

Explicación:

Los dominios que cumplen las restricciones binarias son:

para  $x \geq y+3$     $x \in \{5\}$     $y \in \{2\}$

para  $y \geq z$     $y \in \{2, 4, 6\}$     $z \in \{1, 2, 3\}$

para  $z = x - 3$     $x \in \{5\}$     $z \in \{1, 2\}$

Aplicando AC3: (eliminamos inconsistencias) obtenemos los dominios restringidos:

$CD_x = \{5\}$

$CD_y = \{2\}$

$CD_y = \{1, 2\}$

Siendo las dos soluciones posibles:

$x=5 y=2 z=1$

$x=5 y=2 z=2$

#### T4pregunta082\_1

¿Cuál de las siguientes afirmaciones no es correcta respecto a los métodos de resolución CPS?: {  
    ~ Backtracking construye la solución de manera gradual  
    = Generación y test genera parte de las posibles soluciones al problema  
    ~ Backjumping es parecido a Backtracking pero el retroceso no lo hace de manera gradual.  
}

Explicación: Generación y test genera todas las posibles soluciones y comprueba cuales de ellas satisfacen las restricciones.

#### T4pregunta851\_1

Según el CSP binario:{

- = Todo problema n-ario se puede formular como un problema binario.
- ~ Todo problema binario se puede formular como un problema n-ario.
- ~ Ninguna de las anteriores es correcta.

}

Solución: Transparencia 6 del tema 4.

#### T4pregunta497\_1

El Backjumping:{

- = Es parecido al BT pero el retroceso no se hace a la variable más profunda que está en conflicto con la variable actual.
- ~ Consiste en comprobar comprueba hacia delante todos los valores de las futuras variables que están restringidas con la variable actual.
- ~ Basa su funcionamiento en la programación de restricciones para transformar el problema en otro más sencillo sin inconsistencias de arco.

}

Solución: Transparencia 13 del tema 4.

#### T4pregunta594\_1

El problema de la reinas tiene como características.

{

- ~Dominios continuos y restricciones binarias
- ~Dominios discretos y restricciones múltiples
- =Dominios discretos y restricciones binarias

}

Explicación: El domino del problema son las filas y columnas del tablero (dominio discreto) y las restricciones serás binarias la reina no puede estar en la misma columna o en la misma diagonal, Error corregido

#### T4pregunta19J\_1

¿Cuál de los siguientes afirmaciones sobre "backtracking" no es aplicable?{

- ~ Trashing e inconsistencia de nodo, inconsistencia de arista y dependencia de la ordenación son limitaciones de "backtracking"
- ~ Backtracking construye una solución parcial y extiende esta la solución parcial, incluyendo una variable cada vez hasta llegar una solución total
- = En su forma básica, la idea de backtracking se asemeja a un recorrido en anchura dentro de un grafo no dirigido

}

Explicación:

En su forma básica, la idea de backtracking se asemeja a un recorrido en **profundidad** dentro de un grafo **dirigido**. Cada vértice del grafo es un posible estado de la solución del problema. Cada arco del grafo representa la transición entre dos estados de la solución (i.e., la toma de una decisión)

#### T4pregunta728\_1

¿Cuál de las limitaciones del Backtraking no es solucionada por el Forward Checking? {

- ~ Trashing e inconsistencia de nodo
- ~ Inconsistencia de arista
- = Dependencia de la ordenación

}

Explicación: Al eliminar de sus dominios (temporalmente) los valores de las variables futuras que son inconsistentes con la asignación actual, se evita evaluar aquellas combinaciones que sufren algún tipo de inconsistencia.

Sin embargo, el orden en que se realice la evaluación continúa afectando la velocidad con la cual se encuentra una solución.

#### T4pregunta675\_1

- Cuando se produce la vuelta atrás en el algoritmo de backtracking o backjumping {
- ~ Nunca
  - ~ Cuando llega a la profundidad máxima y encuentra la solución.
  - = Cuando no se puede encontrar una asignación legal de variables que cumpla con las restricciones impuestas.
- }

Explicación:

Batraking: la vuelta atrás se produce cuando no se puede encontrar ninguna asignación legal para una variable.

Backjumping: se produce en las mismas circunstancias, pero con la diferencia de que vuelve atrás directamente a la fuente del problema.

#### T4pregunta013\_1

Si se cumple que **XX = YY - 22** y además que **XX + YY = 176**, que valor tomaría la ecuación:

$$Y(Y - X)X = \{$$

- = 126
  - ~ 123
  - ~ 333
- }

Explicación: Sustituimos en la segunda ecuación la primera

$$YY - 22 + YY = 176 \rightarrow YY + YY = 176 + 22 \rightarrow 2YY = 198 \rightarrow YY = 198/2 = 99$$

$$Y = 9$$

Ahora, sustituimos Y en la primera ecuación

$$XX = 99 - 22 \rightarrow XX = 77$$

$$X = 7$$

Por tanto, el resultado es

$$Y(Y-X)X = 126$$

#### T4pregunta509\_1

En cuanto a los métodos de resolución:

{

- = Una de las limitaciones de Backtracking es la dependencia del orden de selección de las variables. Las fórmulas heurísticas de selección ayudan a seleccionarlas.
- ~ "Generación y Test" genera de forma sistemática y exhaustiva cada una de las posibles asignaciones a las variables retrocediendo a una instancia anterior
- ~ "Generación y Test" no genera todas las asignaciones posibles, construye soluciones posibles y escoge la más óptima

}

*Explicación: Las dos últimas respuestas son incorrectas ya que "Generación y Test" SI genera asignaciones posibles aunque NO escoge la solución más óptima, sino que cuando encuentra una que satisfacen las posibles restricciones, la devuelve. Otro dato es que NUNCA retrocede para encontrar esas asignaciones. Por otra parte, sí que es cierto que la eficiencia del algoritmo Backtracking depende mucho del orden de selección de variables, y se utilizan heurísticas para seleccionarlas.*

#### T4pregunta444\_1

Cuando en AC3 no quedan elementos en el dominio de una variable, se determina que:

{

- ~ Existe una solución al problema.
- ~ Existen múltiples soluciones.
- = Ninguna de las anteriores.

}

*Explicación: Al no existir ningún elemento que satisfaga las restricciones el problema es inconsistente y por lo tanto no tiene solución.*

#### T4pregunta239\_1

En cuanto al Forward checking: {

= Se eliminan los valores de sus dominios, de forma temporal, que son inconsistentes con respecto a la asignación actual.

~ Extiende una solución parcial hasta llegar a la solución total.

~ Transformar el problema en otro más sencillo sin inconsistencias de arco  
}

Explicación: Esto es debido a que FC realiza la búsqueda basándose en su situación actual, por eso ante la previsión de valor inconsistente se elimina del desarrollo de la búsqueda.

#### T4\_pregunta326\_1

¿Qué método de resolución de búsqueda de los elementos enumerados abajo es más apropiado para reducir el espacio durante la búsqueda?

{

= backjumping

~ Backtracking

~ Generación y test

}

explicación:

el retroceso no se hace a la variable

instanciada anteriormente sino a la variable más profunda que está en conflicto con la variable actual.

#### T4pregunta556\_1

En referencia al Forward Checking... {

- ~ Si el dominio de una variable futura se queda vacío, seguiremos explorando las siguientes variables sin ningún tipo de restricción, puesto que esta restricción desaparece.
- = En el caso de obtener un dominio vacío en una variable futura, se deshará la instanciación de la misma y probaremos con un nuevo valor.
- ~ Al obtener un dominio vacío tanto en una variable actual como en una inmediatamente futura, no seguiremos explorando más y diremos a ciencia cierta que no hay valores consistentes.

}

Explicación: La funcionalidad de Forward Checking es precisamente ir comprobando la asignación actual con todos los valores de futuras variables que estén restringidas a dicha variable actual. Por lo que con esta aclaración vemos que la primera y la tercera respuestas, no cumple por completo esto. Mientras que la segunda no nos hace pensar lo contrario.

#### T4pregunta800\_1

Si el dominio de una variable futura se queda vacío (después de la eliminación de los valores de las variables futuras que son inconsistentes con la asignación actual){

- = la asignación actual no es una solución
- ~ está bien porque se reduce la búsqueda
- ~ las restricciones no se incumplen

}

Explicación:

Si el dominio de una variable futura se queda vacío, la asignación actual no es una solución y la instanciación de la variable actual se deshace y se prueba con nuevo valor.

#### T4pregunta379\_1

Un grafo inconsistente:

- {
- = no tiene solución.
- ~ puede tener una o varias soluciones.
- ~ puede tener una única solución, o incluso ninguna.
- }

Explicación: siempre que un grafo sea inconsistente no hay solución, puesto que el dominio es vacío. Pero si el grafo es consistente no implica necesariamente que tenga solución.

#### T4Pregunta895\_1

Cuál de estas afirmaciones sobre Backtracking es CORRECTA {

- ~ Backtracking construye una solución total para luego extenderla para eliminar soluciones y llegar a la correcta.
- = Backtracking construye una solución parcial para luego extenderla hacia una total, llegando así a la solución.
- ~ Backtracking construye una solución parcial eliminando futuras decisiones para acortar el camino hacia la solución.

}

Explicación: Backtracking construye una solución parcial satisfaciendo las restricciones, luego la extiende incluyendo la variable hasta llegar a la solución.

#### T4pregunta435\_1

En el algoritmo AC3 de búsqueda por CSP, la variable Q:

$$Q = \{c(e_p) = \langle V_i, V_j \rangle | e_p \in E, i \neq j\}$$

Mientras  $Q \neq \emptyset$  hacer

$\langle V_k, V_m \rangle = \text{seleccionar\_y\_borrar}(Q)$

cambio = falso

Para todo  $v_k \in D_k$  hacer

Si no\_consistente ( $v_k, D_m$ ) entonces

borrar ( $v_k, D_k$ )

cambio = cierto

FinSi

FinPara

Si  $D_k = \emptyset$  entonces salir\_sin\_solución FinSi

Si cambio = cierto entonces

$Q = Q \cup \{c(e_r) = \langle V_i, V_k \rangle | e_r \in E, i \neq k, i \neq m\}$

FinSi

FinMientras

{

= contiene todas las restricciones binarias del problema en ambos sentidos.

~ abarca todas las restricciones binarias del problema partiendo de 0.

~ ninguna de las anteriores.

}

Explicación: el algoritmo AC3 crea un grafo que une cada par de variable-dominio.

Una vez creado estudia todos los conectores para poder decidir cuales no cumplen con las restricciones dadas. Por ello, ha de contener todas las posibilidades en ambos sentidos, para ir eliminando las que no cumplan con dichas restricciones.

#### T4pregunta467\_1

¿Qué diferencia a backjumping de backtracking?

{

= Backjumping realiza el retroceso a la variable más profunda que está en conflicto con la actual.

~ Backjumping realiza el retroceso a la variable instanciada anteriormente.

~ Backjumping realiza búsqueda en profundidad.

}

Explicación: Las respuestas 2 y 3 son características de backtracking.

#### T4pregunta312\_1

En los métodos de resolución por búsqueda: ¿Cuándo se genera de forma sistemática y exhaustiva cada una de las posibles asignaciones a las variables y comprobar si satisfacen todas las restricciones. Se explora el espacio definido por el producto cartesiano de los dominios de las variables?:{

=Generación y test.

~Backtracking.

~Backjumping.

}

Explicación:

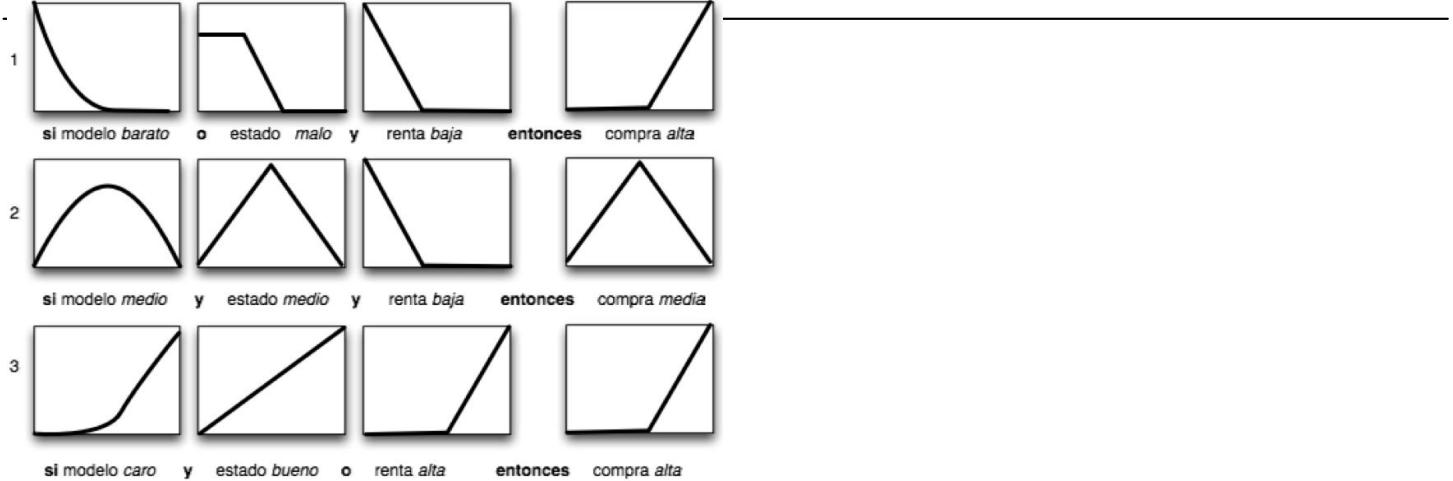
Generación y test : generar de forma sistemática y exhaustiva cada una de las posibles asignaciones a las variables y comprobar si satisfacen todas las restricciones. Hay que explorar el espacio definido por el producto cartesiano de los dominios de las variables.

Backtracking : se trata de construir la solución de forma gradual, instanciando variables en el orden definido por la permutación dada.

Backjumping: parecido al Backtracking pero el retroceso no se hace a la variable instanciada anteriormente sino a la variable más profunda que está en conflicto con la variable actual.

# Preguntas SI Foro 5

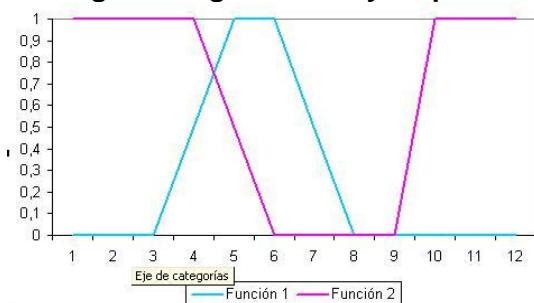
1. Un modificador lingüístico permite modificar el significado de un conjunto difuso. Por ejemplo:  
 $= u_{\text{poco}}(x) = x^y$ ,  $y < 0$  es un modificador que aplica el significado muy poco a un conjunto.
2. Tenemos el siguiente sistema experto difuso que permite obtener la posibilidad de venta de un coche de segunda mano dada la renta del cliente. Asumiendo una escala de valoración de 0 a 10 deseamos saber si dado un modelo de coche valorado con 9 y con estado 7 se venderá para clientes con renta alta (8), asumiendo los operadores por defecto.



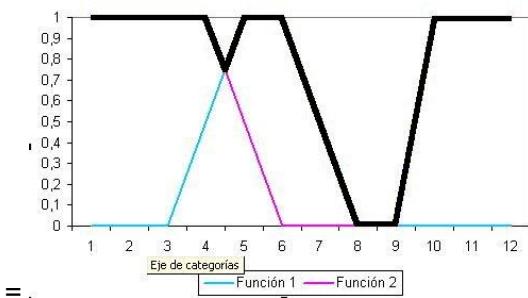
De esta manera, el resultado del sistema experto una vez agregado el resultado de las reglas es, aproximadamente, el siguiente conjunto difuso:{



3. Según la lógica difusa y disponiendo de estas dos funciones:



¿Cuál representa la unión?



4. En los sistemas expertos difusos es cierto que:

= Se pueden combinar varias reglas mediante varios posibles métodos de agregación

**5. Existen varias formas para crear programas que actúen como sistemas expertos:**

= Los primeros y más utilizados son los sistemas basados en reglas.

**6. Las características esenciales de un sistema experto son:**

= Alto desempeño, Tiempo de respuesta adecuado, Confiabilidad, Comprensible, Flexibilidad y que tenga una representación explícita del conocimiento.

**7. Los parámetros a establecer en el SE (Sistemas expertos Difusos) son: {**

= El And/Or a utilizar, el método de agragación para los conjuntos de variables a defuzzificar, el método de activación y el método de defuzzificación.

**8. ¿Cuáles de las siguientes sentencias definen el término 'Variable Lingüística'?:**

= Son aquellas palabras o sentencias que se van a enmarcar en un lenguaje predeterminado.

**9. El orden a aplicar las reglas de un sistema difuso es:**

= Fuzzyficación, aplicar operador fuzzy, implicador fuzzy, combinación de las reglas, defuzzificación.

**10. ¿Cuál o cuáles de las siguientes variables serían aptas como variables lingüísticas?**

~ Edad.

~ Distancia.

= Las dos anteriores son correctas.

**11. ¿Cuál de las siguientes sentencias es correcta?:**

= La función de pertenencia se establece de una manera arbitraria.

**12. ¿Qué son los Modificadores lingüísticos en lógica difusa?:**

= Operador que modifica el significado de un conjunto difuso.

**13. Conseguimos una representación del conocimiento de forma más natural mediante:**

= Lógica difusa

**14. Un sistema experto:**

= Representa y usa conocimiento y puede operar con información incompleta.

**15. ¿Cuál de las siguientes respuestas no es una alternativa a la incertidumbre?:**

= Teorema de Bolzano.

**16. En los sistemas expertos difusos:**

~ Se genera un consecuente para cada regla en función del grado de cumplimiento de cada una.

~ Se obtiene una salida numérica a partir de todos los consecuentes obtenidos.

= Ambas son correctas.

**17. Tres de las características esenciales de los sistemas expertos son:**

~ Bajo desempeño, Confiabilidad y Flexibilidad.

~ Alto desempeño, Tiempo de respuesta adecuada y Representación implícita del conocimiento.

= Ninguna de las anteriores.

**18. ¿Qué es la defuzzificación?:**

= Se obtiene una salida numérica a partir de todos los consecuentes obtenidos.

**19. El alto desempeño de un Sistema Experto implica:**

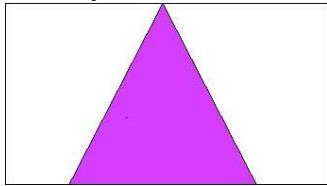
= La capacidad de mantenerse activo y en funcionamiento durante largos períodos de tiempo, sin descanso.

**20. En los sistemas expertos el conocimiento se representa:**

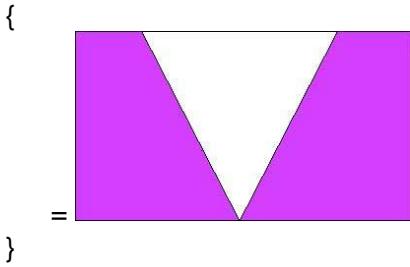
= Mediante lógica (proposiciones y predicados)

- 21. Si tuviéramos un sistema experto con 3 reglas y obtenemos un 0% de pertenencia a la regla 1, un 40% para la regla 2 y un 80% para la 3, actuaríamos de la siguiente manera:**  
= Se aplican todas las reglas, pero en función del grado de cumplimiento de cada una.
- 22. ¿Qué realiza la parte de 'toma de decisiones' en un Sistema experto Difuso:**  
= Se comparan los valores de la entrada al sistema con las funciones asociadas a la regla de cada entrada
- 23. ¿Qué es el COG?**  
= Un método de defuzzificación.
- 24. En los conjuntos difusos es FALSO que:**  
= Solo está definida la operación AND.
- 25. ¿Cuál de estas características no pertenecen normalmente a los Sistemas Expertos?:**  
= Representa y usa datos.
- 26. Los sistemas expertos son una rama de la IA que:**  
= hace uso del conocimiento especializado para resolver problemas como un especialista humano.
- 27. La suma de los factores de pertenencia de un conjunto difuso para un determinado valor puede ser:**  
= Cualquier valor.
- 28. En la fase de "defuzzyificación", es correcto que:**  
= Podemos utilizar el cálculo de centro de masas para la obtención del resultado.
- 29. Diferencias entre Sistema Clásico y Sistema Experto:**  
= El sistema clásico representa y usa datos frente al experto que representa y usa conocimiento.
- 30. En lógica difusa, un modificador lingüístico:**  
= Es un operador que modifica el significado de un conjunto difuso.
- 31. Las siglas FCL pertenecen a:**  
= Fuzzy Control Language
- 32. Los términos probabilidad y pertenencia..:**  
= Se diferencian en que el conjunto de probabilidades siempre suma uno y el de la pertenencia no necesariamente suma uno.
- 33. ¿En qué tipo de lógica se establece una función de pertenencia?**  
= Lógica Difusa
- 34. En los sistemas expertos, ¿suelen darse situaciones de incertidumbre?**  
= Sí, pueden darse.
- 35. En un conjunto difuso:**  
= No se especifican los elementos que forman parte del conjunto dependiendo de si cumplen o no unas propiedades sino que se especifica una función de pertenencia que indica si un elemento pertenece al conjunto dado.

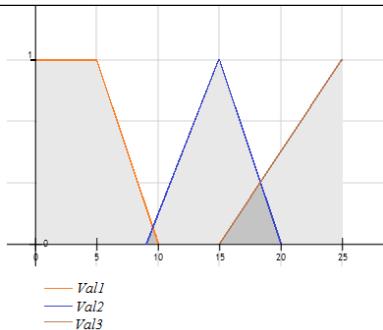
36. Representando la zona coloreada de la siguiente figura un conjunto difuso  $\mu_A(x)$ .



Cuál de los siguientes conjuntos (zona coloreada del dibujo) corresponde a  $\mu_{\neg A}(x)$ .



37. ¿Cómo inicializarías la variable lingüística 'G' (FCL) con los conjuntos val1, val2 y val3?



a)

FUZZIFY G

```
TERM val1 := (0,1) (5,1) (10,0);  
TERM val2 := (9,0) (15, 1) (20, 0);  
TERM val3 := (15, 0) (25, 1);  
END_FUZZIFY
```

38. En la toma de decisiones de un Sistema Experto, las reglas se ejecutan:

= De forma paralela.

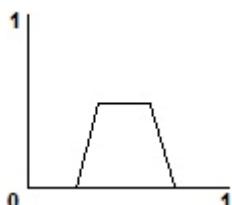
39. Los sistemas basados en lógica difusa intentan imitar la forma:

= del racionamiento de los humanos.

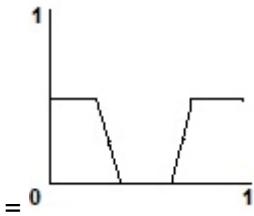
40. Si disponemos de una impresora con lógica difusa que clasifica los colores según su pertenencia a los valores lingüísticos blanco, negro y gris. Si nos dice que el color X tiene una pertenencia 0.3 al blanco y 0.4 al gris, podemos asegurar que:

- ~ Tendrá una pertenencia de 0.3 al negro.
- ~ Que el color será 30% blanco.
- ~ Las dos anteriores son falsas.

41. Tenemos el siguiente conjunto:



Cuál es el conjunto resultante al aplicar la ecuación  $\mu_{\neg A}(x) = 1 - \mu_A(x)$



}

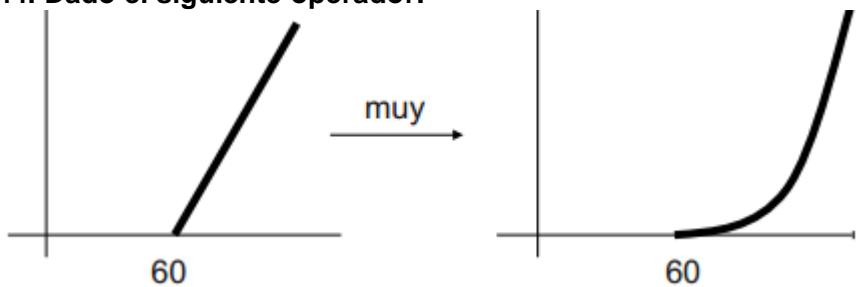
**42. Podemos afirmar que en los sistemas expertos difusos se da la acción de:**

- = Comparar los valores numéricos de entrada al sistema con las funciones de pertenencia asociadas a los términos lingüísticos de la parte del antecedente de la regla asociada a esa entrada.

**43. Con respecto a los modificadores lingüísticos como "Muy..." o "Poco...", indica cuál de estas afirmaciones es correcta:**

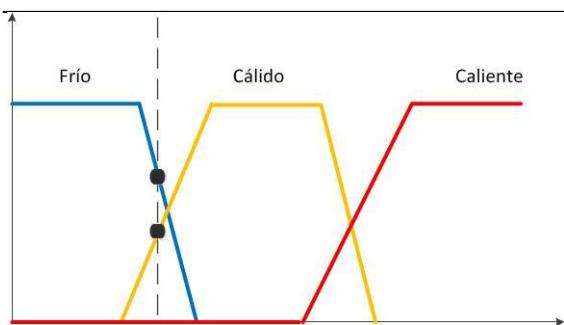
= Muy  $\mu(x) = \mu(x)^2$

**44. Dado el siguiente operador:**



**En Lógica Difusa, el modificador lingüístico Muy, ¿cómo lo podríamos describir?**

= Muy  $\mu(x) = \mu(x)^2$



**45. ¿Qué valor lingüístico no tiene la línea vertical?**

= caliente

**46. ¿En qué se diferencian la Lógica de primer orden y la Lógica difusa?:{**

- = La lógica difusa tiene más facilidad para representar el conocimiento real que la lógica de primer orden.
- }

**47. Indica a qué nivel permiten representar la información los Sistemas Expertos: {**

- = A nivel de los seres humanos.
- }

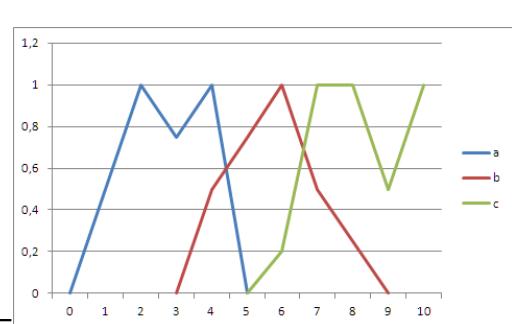
**48. ¿Cuál de las siguientes variables prodrían ser variables lingüísticas?{**

- ~ Temperatura.
  - ~ Edad.
  - = Las dos son correctas.
- }

**49. Dada la siguiente inicialización de la variable SI: FUZZIFY SI**

```
TERM a := (0, 0) (2, 1) (3, 0.75) (4, 1) (5, 0);  
TERM b := (3, 0) (4, 0.5) (5, 0.75) (6, 1) (7, 0.5) (9, 0);  
TERM c := (5, 0) (6, 0.2) (7, 1) (8, 1) (9, 0.5) (10, 1);  
END_FUZZIFY
```

**¿Qué conjunto de los siguientes podemos indicar que representa a la variable SI? {**



**50. ¿Cuáles son los inconvenientes de la lógica difusa?{**

- ~ Es monotónica
  - = Ninguna de las anteriores es correcta
  - ~ La dificultad que encontramos al intentar representar el conocimiento real
- }

**51. ¿Qué tipo de lógica representa de una forma más natural el conocimiento real? {**

- = Lógica difusa
- }

**52. Dados los siguientes conjuntos:**

**DEFUZZIFY tip**

```
TERM cheap := (0,0) (5,1) (10,0);  
TERM average := (10,0) (15,1) (20,0);  
TERM generous := (20,0) (25,1) (30,0);
```

**¿Qué ocurriría si tip=10?**

- = El centro de masas valdría 0.

**53. ¿Cuál de las siguientes afirmaciones es correcta?{**

- ~ Los SE necesitan representar y usar datos.
- ~ Los SE representan de forma implícita el conocimiento.
- = Ninguna de las anteriores es correcta.

**54. Pertenencia y probabilidad:** {

= No podemos establecer el valor de los conjuntos de pertenencia sin conocer su función correspondiente.

**55. Respecto a las diferencias entre un sistema experto y uno clásico, podemos afirmar que:** {

= El sistema experto puede funcionar con pocas reglas.

}

**56. En Fuzzy Control Lenguaje (FCL) la declaración de las variables de entrada se hace en el apartado:**

{  
    =VAR\_INPUT  
}

**57. En cuanto a las operaciones entre conjuntos, dada la siguiente fotografía...**

$$\mu_{A \cup B}(x) = \min(\mu_A(x), \mu_B(x))$$

$$\mu_{A \cap B}(x) = \max(\mu_A(x), \mu_B(x))$$

{  
    ~ La función min. es incorrecta.  
    ~ La función max. es incorrecta.  
    = Ambas son incorrectas.  
}

**58. Fuzzy Control Language:**

= No es un lenguaje de lógica difusa totalmente completo.

**59. En cuanto a la sintaxis para definir reglas de un algoritmo fuzzy utilizando la FCL, es incorrecta:**

= RULE 1: THEN condition IF conclusion;

**60. Respecto a Fuzzy Control Language {**

= Es un lenguaje que sólo tiene características propias de lógica difusa.

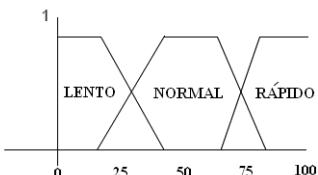
}

**61. ¿Cuál de los siguientes operaciones entre conjuntos de logica difusa corresponde a la operacion booleana**

"AND" (x AND y)?

= minimum(truth(x), truth(y))

**62. Utilizando la siguiente gráfica, velocidad de un auto**



**Si el auto es rápido, es posible afirmar**

**que:**

= Su pertenencia a una velocidad "lenta" es menor o igual a 0

**63. Cuáles de las siguientes son características esenciales de un Sistema Experto:**

= Confiable, comprensible y flexible.

}

**64. Si un conjunto difuso A de dominio D, viene caracterizado por una función de pertenencia  $f_A(x)$  que asocia a cada elemento x del dominio, un valor en el intervalo [0,1] que determina su grado de pertenencia a ese conjunto. Elige la expresión matemática correcta:{**

= la función de pertenencia,  $f_A(x) \in [0, 1] \forall x \in D$

**65. Las características de un sistema experto son:{**

= Alto desempeño, tiempo de respuesta adecuado, confiabilidad, comprensible, flexibilidad y representación explícita del conocimiento.

**66. ¿Qué es la Lógica Difusa? {**

=Es un tipo de lógica que reconoce valores que representan grados de veracidad o falsedad.

}

**67. El operador de conjunto difuso unión,  $A \cup B : f_{A \cup B}(x) = \max[f_A(x), f_B(x)]$ , cumple la propiedad: {**

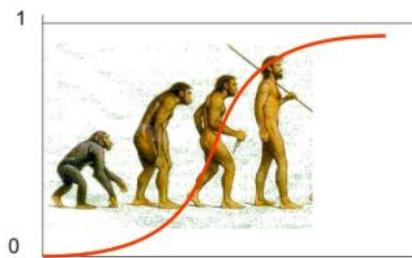
= De Morgan  $\neg(A \cup B) = \neg A \cap \neg B$

}

**68. Indica la respuesta correcta:**

$$= A \cup (\neg B \cap C) = (A \cup \neg B) \cap (A \cup C)$$

**69. ¿A qué pertenece la siguiente imagen?**



= Lógica difusa.

**70. La representación del conocimiento de una forma más natural se le atribuye a:**

= Lógica difusa

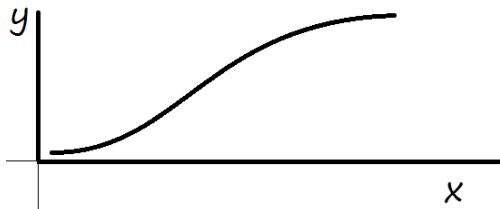
**71. La agregación en un sistema difuso es:**

= El proceso que genera el conjunto de salida a cierta variable

**72. Cual de los siguientes elementos es una ventaja de un sistema experto sobre lo sistema clásico:{**

- = El sistema puede funcionar con poca reglas

**73. Dada la grafica anterior podriamos afirmar que se podria tratar de :{**



- = Logica difusa.

**74. De que partes consta un sistema experto difuso:**

- = De una entrada de datos, una base de conocimientos, una fuzzyfication, un toma de decisiones, una defuzzyfication.

**75. ¿Cuál de las siguientes lógica no es monotónica y no tiene dificultades para representar el conocimiento real?{**

- = Lógica difusa

}

**76. ¿Cuál de los siguientes algoritmos es incorrecta?{**

- = ASUM:  $u_1(x) - u_2(x) + u_1(x) u_2(x)$

**77. En cuanto a la lógica difusa:**

- = Intenta disminuir las transacciones entre estados combinando reglas para decidir entre conjunto de estos aplicando el cumplimiento de dichas reglas y generando soluciones para cada una de ellas. Dependiendo del grado de cumplimiento se escoge el resultado de una de ellas.

**78. Al enmarcar una variable lingüística ¿Que debemos tener en cuenta?**

- = Tendremos en cuenta tanto el universo como los valores lingüísticos

**79. Indica cuál es desventaja de la lógica difusa: {**

- ~ Tiene múltiples definiciones de operadores y reglas de inferencia difusa.
- ~ Ante un problema que tiene solución mediante un modelo matemático, obtenemos peores resultados usando lógica difusa.
- = Todas las anteriores son correctas.

}

**80. Acerca de sistemas expertos, es verdad:**

- = Debe ser capaz de explicar paso a paso cómo se obtuvo la respuesta

**81. ¿Cuál de las siguientes características es una característica propia de la lógica difusa?**

- = Dice cuanto porcentaje de cada conjunto le pertenece a una variable.

**82. Según la definición de variable lingüística podemos decir:**

- = Es una variable cuyos valores son palabras o sentencias en un lenguaje natural o sintético.

**83. Pregunta:** "no", "muy", "algo", "casi" serian ejemplos de... {

= Modificadores lingüísticos

**84. Cuál de las siguientes afirmaciones es verdadera:**

= Los sistemas difusos nacen por la necesidad de interpretar acciones en el mundo real.

**85. Cuando aplicamos incertidumbre (propiedad de modularidad) en un ejemplo como:**

- Si A entonces B con probabilidad 0'7
- Si C entonces B con probabilidad 0'8

En el caso de que se tenemos A y C, ¿cuál es la probabilidad de B? :{

= B no se puede obtener.

**86. ¿Qué forma pueden adoptar las funciones de pertenencia?:{**

- ~ Trapezoidal
- ~ Campana de Gauss
- = Las tres respuestas son correctas

**87. Dados la siguiente sentencia que representa una parte de código de un programa FCL: TERM X:= (5,1);**

**¿Qué representa?**

= Para un valor de "5", representando "X", su pertenencia es "1".

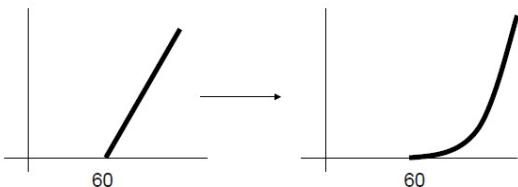
**88. En lógica difusa, más, menos y muy, son:**

= Modificadores lingüísticos

**89. Una variable lingüística puede:{**

- ~ Tras ser aplicada a un universo de discurso generar un universo reducido (sub universo)
  - ~ Ser aplicada conjuntamente con un modificador lingüístico
  - = Las dos respuestas son ciertas
- }

**90. ¿Cuál de los siguientes operadores se ha aplicado en la imagen anterior? {**



= Muy

**91. Con estas reglas, qué propina se daría (mucha, normal o poca) si la comida fue normal y el servicio bueno:**

RULE 1: IF servicio IS pobre OR comida IS mala THEN propina IS poca;

RULE 2: IF servicio IS normal AND comida IS buena THEN propina IS mucha;

RULE 3: IF servicio IS normal OR comida IS normal THEN propina IS normal;

RULE 4: IF servicio IS bueno AND comida IS normal THEN propina IS mucha;

{

= normal

}

**92. Respecto a las partes de un sistema difuso experto, indica cual de las siguientes afirmaciones es correcta respecto a la "fuzzyficación":{**

- ~Los antecedentes de las reglas difusas no tienen por qué cumplirse siempre.
- ~Se disparan las reglas cuyo consecuente tiene un cierto grado de cumplimiento.
- =Ninguna de las anteriores es correcta.

}

**93. Según las partes de los Sistemas expertos Difusos, se obtiene:**

- = Una salida numérica a partir de todos los consecuentes obtenidos.

**94. A partir de estas reglas, razona cuál sería la nota (mala, regular, buena) si la caligrafía es normal y el nº de faltas medio:**

- Regla 1: Si la caligrafía es mala o hay muchas faltas. La nota es mala.
- Regla 2: Si la caligrafía es normal y hay un nº medio de faltas. La nota es regular.
- Regla 3: Si la caligrafía es normal y no hay faltas. La nota es buena
- Regla 4: Si la caligrafía es buena o el nº de faltas es medio. La nota es buena.

- = La nota es buena debido a la agregación.

## TEMA 6 ÁRBOLES DE DECISIÓN

**Ganancia de información.** Qué expresión correcta, corresponde a la ganancia de y condicionado a x: {

=  $IG(Y | X) = E(Y) - E(Y | X)$ .

~  $IG(Y | X) = E(Y) - E(X)$ .

~  $IG(Y | X) = E(X) + E(Y | X)$ .

}

*Explicación:*  $IG(Y | X) = E(Y) - E(Y | X)$ , la ganancia implica que el atributo X permite reducir la incertidumbre de la clasificación del ejemplo de entrada.

**Entropía:** Hablamos de "DISTRIBUCIÓN EN PICO" cuando se cumple que:{

=  $P_i = 1$  y  $P_j=0$ , para todo  $j \neq i$

~  $P_i = 0$  y  $P_j=1$ , para todo  $j \neq i$

~  $P_i=1$  y  $P_j=0$ , para todo  $j=i$

}

*Explicación:* Como bien viene indicado en el tema 6 página 5, la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información. O lo que es lo mismo  $P_i = 1$  y  $P_j=0$ , para todo  $j \neq i$

**¿Qué es la entropía?:{**

~ Es la medida del grado de incertidumbre asociado a una distribución uniforme.

~ Es la medida del grado de incertidumbre siempre que consiga la máxima información.

= Es la medida del grado de incertidumbre asociado a una distribución de probabilidad.

}

*Explicación:* La tercera opción es la correcta. La distribución uniforme o de pico (en esta última es en la que conseguimos máxima información) son tipos de distribuciones. Sin embargo, la entropía es una medida y trabaja con distribuciones de probabilidad.

### **Los árboles de decisión {**

- ~ Es interesante aprenderlos a partir de un grafo
  - ~ Establecen el vector de entrada.
  - = Ninguna de las anteriores
- }

Explicación: La solución es la c) porque es interesante aprenderlos a través de un conjunto de vectores(no grafos) y establecen el orden para testar los atributos y conseguir la clasificación del vector de entrada.

### **En cuanto a la entropía:{**

- ~ Es mínima cuando la distribución es uniforme.
  - ~ Es máxima cuando la distribución es de pico.
  - = Es máxima cuando la distribución es uniforme.
- }

Respuesta: en una distribución uniforme, todos los valores son igualmente probables  $P_i=1/N$  y por tanto la entropía es máxima, lo cual indica máxima incertidumbre (transparencia 5)

### **Los arboles de decisión establecen {**

- = *en qué orden testar los atributos para conseguir la clasificación del vector de entrada.*
  - ~ *a qué velocidad testar los atributos para conseguir la clasificación del vector de salida.*
  - ~ *el estado inicial de la clasificación de los nodos.*
- }

Explicación: Según los apuntes, en la página 3 del Tema 6. Una de las características de los árboles de decisión es que: establece en qué orden testar los atributos para conseguir la clasificación del vector de entrada

**¿Qué es la ganancia de información? {**

- ~ Conjunto de todos los valores posibles de X
  - = Medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y.
  - ~ Ninguna de las anteriores
- }

Explicación: La ganancia de información es la medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y.

**Sobre atributos con gran número de valores: {**

- ~ Se forman grupos pequeños de ejemplos que siempre deben ser homogéneos.
- = Se forman grupos pequeños de ejemplos que pueden ser homogéneos por casualidad.
- ~ Ningunas de las anteriores.

}

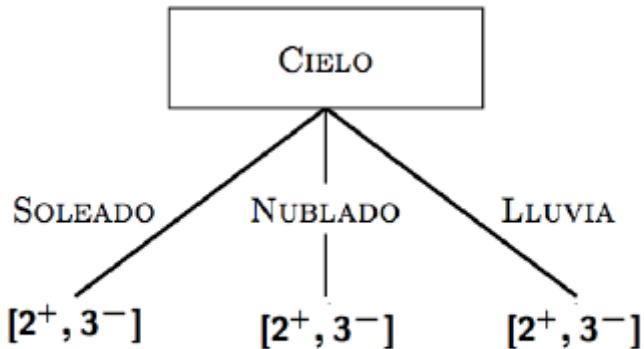
Explicación: Transparencia 11 del Tema 6. Se observa que en el apartado de atributos con gran número de valores dice que se forman grupos pequeños de ejemplos que pueden ser homogéneos por casualidad.

**Cuando el conocimiento de X mejora la información que se dispone sobre Y, se dice que tiene un entropía condicionada :{**

- = menor que  $E(Y)$ .
  - ~ igual que  $E(Y)$ .
  - ~ mayor que  $E(Y)$ .
- }

Explicación: en el tema 6 diapositiva 6. Explica que una entropía condicionada menor que  $E(Y)$  indica que el conocimiento de X mejora la información que se dispone sobre Y.

¿Cuál será la entropía del siguiente atributo?



{

= 1.585

~ Para atributos uniformemente distribuidos la entropía será máxima, es decir, 1

~ Para atributos uniformemente distribuidos la entropía será mínima, es decir, 0

}

Explicación: En este caso estamos en una situación de una variable uniformemente distribuida por lo que su entropía será máxima. Sin embargo, según la ecuación para calcular la entropía:  $E(\text{Cielo}) = -(1/3)\log_2(1/3) - (1/3)\log_2(1/3) - (1/3)\log_2(1/3) = 1.585$

¿ID3 con qué tipo de atributos trabaja? {

~ Atributos discretos

= Todas las anteriores

~ Atributos continuos si están descompuestos en rangos. }

Explicación: Ambas respuestas son correctas, ID3 trabaja con atributos discretos y, si utilizamos continuos hay que descomponerlos en rangos ordenando los ejemplos según el valor y tomando como límites los puntos medios de aquellos en que se cambie la clase.

**En una distribución pico en la que  $P_i = 1$  y  $P_j=0$ , para todo  $j \neq i$  la entropía será{**

- ~ Máxima
- = Mínima
- ~ No lo podemos saber }

*Explicación: Tal y como aparece en la transparencia número 5 del tema, bajo estas condiciones, la entropía es mínima, lo cual indica mínima incertidumbre o sea máxima información.*

**¿Qué implica una alta ganancia de información? {**

- ~ Que el atributo X permite aumentar la incertidumbre de la clasificación del ejemplo de entrada.
  - = Que el atributo X permite reducir la incertidumbre de la clasificación del ejemplo de entrada.
  - ~ Que la clase Y permite reducir la incertidumbre de la clasificación del ejemplo de entrada.
- }

*Explicación: Por definición, una alta ganancia implica que el atributo X permite reducir la incertidumbre de la clasificación del ejemplo de entrada.*

**¿Cual de las siguientes afirmaciones es cierta acerca de la entropía?:{**

- = En una distribución uniforme, todos los valores son equiprobables, lo que implica entropía máxima y por tanto, total incertidumbre.
  - ~ En una distribución pico en la que  $P_i = 0$  y  $P_j = 1$  para  $j \neq i$  la entropía es mínima.
  - ~ En una distribución uniforme todos los valores son equiprobables, lo que implica una entropía mínima y por tanto, la mínima incertidumbre.
- }

*Explicación: Las distribuciones uniformes deben su nombre a la equiprobabilidad de sus valores, y por definición, esto implica la máxima entropía y máxima incertidumbre.*

*Por otro lado, las distribuciones pico cumplen la definición si  $P_i = 1$  y  $P_j = 0$ .*

**El recorrido de un árbol de decisión equivale a una:{**

- = Búsqueda irrevocable.
- ~ Búsqueda tentativa.
- ~ Ninguna de las anteriores.

} Explicación: Los árboles de decisiones se recorren mediante un método voraz, por lo tanto irrevocable

**Una de las características de los árboles de decisión es: {**

- = Establece en qué orden testar los atributos para conseguir la clasificación del vector de entrada.
  - ~ En una distribución uniforme, todos los valores son igualmente probables.
  - ~ Ninguna de las anteriores.
- }

Explicación: Características de arboles de decisión:

- Estructura para clasificación de vectores de atributos.
- Establece en qué orden testar los atributos para conseguir la clasificación del vector de entrada.
- Para componer dicho orden se eligen primero aquellos atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.
- Es interesante aprenderlos a partir de un conjunto de vectores

**En la entropía, ¿qué grado de incertidumbre tendremos en una distribución uniforme? {**

- ~ Mínima
  - = Máxima.
  - ~ No se puede saber
- }

Explicación: Como hemos visto en clase, la entropía no es una probabilidad, es una incertidumbre; y en este caso, la incertidumbre es máxima.

**El algoritmo ID3, a la hora de descomponer los atributos continuos en rangos, después de ordenarlos: {**

- = Debemos tener en cuenta los puntos límites donde se cambie de clase.
  - ~ No es necesario descomponer, pues siempre tratamos con atributos discretos.
  - ~ Debemos de realizar la descomposición siempre en partes iguales.
- }

Explicación: En primer lugar realizamos la ordenación de los ejemplos, y luego, realizamos la descomposición tomando como referencia los puntos medios de aquellos que se cambien de clase.

**Lanzamos un dado al aire, ¿cuál es la entropía del lanzamiento?:{**

- 1, máxima incertidumbre
  - 0,16, al ser distribución uniforme
  - = 2,58, máxima incertidumbre.
- }

Explicación: la entropía se calcularía así:  $6 * (-1/6) * \log(1/6) / \log 2$

Es distribución uniforme debido a que tenemos la misma probabilidad de que nos toque una de las 6 caras, por eso multiplicamos por 6.

**Una entropía condicionada menor que E(Y) indica que: {**

- ~ El conocimiento de X reduce la información que se dispone sobre Y.
  - = El conocimiento de X mejora la información que se dispone sobre Y.
  - ~ No indica nada. Siempre es mayor o igual que E(Y).
- }

Explicación: La entropía condicionada puede ser menor que E(Y), lo cual indica que hay una mejora en la información que se tiene de Y.

**En la Entropía condicionada: {**

- ~ No interviene ninguna variable.
  - ~ Interviene una variable.
  - = Intervienen al menos dos variables.
- }

Explicación: La entropía condicionada es una extensión del concepto de entropía de la información a procesos donde intervienen varias variables aleatorias no necesariamente independientes.

### **La Entropía condicionada es según su definición :{**

- ~ la entropía de la distribución de Y condicionada a Y.
  - = la entropía de la distribución de Y condicionada a X.
  - ~ la entropía de la distribución de X condicionada a X.
- }

Explicación: en el tema 6 diapositiva 6. Nos dice en la definición de entropía condicionada que es la entropía de la distribución de Y condicionada a X.

### **¿Qué pasa si usamos ID3 con atributos continuos?{**

- ~ Nada porque es con lo que trabaja.
  - = Hay que descomponerlo en rangos.
  - ~ Hay que ordenarlos según valor y pasarlo a discretos.
- }

Explicación: ID3 solo trabaja con atributos discretos, para el caso de utilizar atributos continuos hay que descomponerlos en rangos. La primera esta clara que esta mal, pero la tercera esta mal porque dice "Hay que ordenarlos según valor" (hasta ahí esta bien) "y pasarlo a discretos" (esto ya esta mal).

### **Los árboles de decisión, ¿contienen información local? {**

- ~ No, no están basados en información local.
  - = Sí, lo que permite poder encontrar la mejor solución.
  - ~ Si, pero nunca encontrarán la mejor solución.
- }

*Explicación: Los árboles de decisión usan una estrategia irrevocable de descenso por gradiente, lo que quiere decir que poseen conocimiento local y buscan la mejor solución a partir de esa información local.*

### **El valor de entropía, ¿está acotado de 0 a 1?: {**

- ~ Si, dado que es un valor de probabilidad.
- ~ Si, dado que es una medida del grado de incertidumbre.

= No, dado que es una medida del grado de incertidumbre.

}

Explicación: Al no ser un valor de probabilidad, cuanto mas cercano a 0 sea el valor menos incertidumbre existe pero el valor puede ser mayor que 1, es decir, 1 no quiere decir que el valor de incertidumbre sea máximo.

**Si tenemos una  $E(Y) = 0.971$  y tras analizar la información de la que disponemos obtenemos los siguientes datos para aplicar una entropía condicionada:**

$v_j$	$\text{Prob}(X=v_j)$	$E(Y X=v_j)$
Atributo 1	0.3	0.92
Atributo 2	0.4	0.81
Atributo 3	0.3	0.92

¿Después de calcular el valor de  $E(Y | X)$  podemos decir que hemos obtenido ganancia de información?

{

= Si, hemos obtenido una ganancia  $\approx 0.09$

~ Si, hemos obtenido una ganancia  $\approx 0.5$

~ No

}

**Explicación:** He utilizado los datos de la diapositiva 9, por si no queda clara la explicación.

\* Primero aplicamos la formula de la entropía condicionada a los datos de la tabla.

$$E(Y | X) = \sum_j \text{Prob}(X=v_j) E(Y | X=v_j)$$

$$E(Y | X) = 0.3 * 0.92 + 0.4 * 0.81 + 0.3 * 0.92 = 0.876$$

\* Una vez hecho esto aplicamos la de la ganancia

$$IG(Y | X) = E(Y) - E(Y | X)$$

$$IG(Y | X) = 0.971 - 0.876 = \textcolor{red}{0.095 \approx 0.09}$$

**La fórmula de la entropía condicionada es:{**

$$= E(Y | X) = \sum_j \text{Prob}(X=v_j) E(Y | X=v_j)$$

$$\sim E(X | Y) = \sum_j \text{Prob}(X=v_j) E(Y | X=v_j)$$

$$\sim E(Y | X) = \sum_j \text{Prob}(Y=v_j) E(Y | X=v_j)$$

}

Explicación: La fórmula correcta para el cálculo de la entropía condicionada, como pone en la transparencia 6 del tema 6, es  $E(Y | X) = \sum_j \text{Prob}(X=v_j) E(Y | X=v_j)$

**Dado un conjunto de ejemplos y su árbol de decisión generado por ID3. Los elementos de dicho árbol son:{**

~ Nodos que contienen valores del padre, arcos que contienen variables continuas y hojas que clasifican el ejemplo como positivo.

~ Nodos que contienen valores del padre, arcos que contienen valores del hijo y las hojas que clasifican el ejemplo como positivo.

= Nodos que contienen atributos, arcos que contienen posibles valores del padre y las hojas que clasifican el ejemplo como positivo o negativo.

}

Explicación: Cada nodo del árbol representa cada uno de los atributos (Antigüedad, Moroso, Ingresos). Los arcos representan los valores que el nodo puede tomar (<600, 600-1200, >1200) y las hojas representan la decisión final que pueden ser positivos o negativos (Conceder, No conceder).

**En los árboles de decisión, que criterio utilizamos para el orden en el que testaremos los atributos: {**

~El que nos da una ganancia de información más baja.

~Es indiferente, el orden es aleatorio.

=El que nos da una ganancia de información más alta.

}

Explicación: Para componer dicho orden se eligen primero aquellos atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.

EJ.	CIELO	HUMEDAD	JUGARTENIS
D <sub>1</sub>	SOLEADO	ALTA	-
D <sub>2</sub>	SOLEADO	ALTA	-
D <sub>3</sub>	NUBLADO	ALTA	+
D <sub>4</sub>	LLUVIA	ALTA	+
D <sub>5</sub>	LLUVIA	NORMAL	+
D <sub>6</sub>	LLUVIA	NORMAL	-
D <sub>7</sub>	NUBLADO	NORMAL	+
D <sub>8</sub>	SOLEADO	ALTA	-
D <sub>9</sub>	SOLEADO	NORMAL	+
D <sub>10</sub>	LLUVIA	NORMAL	+
D <sub>11</sub>	SOLEADO	NORMAL	+
D <sub>12</sub>	NUBLADO	ALTA	+
D <sub>13</sub>	NUBLADO	NORMAL	+
D <sub>14</sub>	LLUVIA	ALTA	-

**Dado el conjunto anterior, que atributo cogeríamos primero para aprender el concepto “días que se juega a tenis” y obtener el nodo inicial del árbol de decisión mediante el algoritmo ID3{**

- ~ El orden en que cojamos los atributos no tiene importancia, el nodo inicial puede ser tanto “cielo” como “humedad”.
- = Cogemos el atributo “cielo”, ya que es el que mayor ganancia de información nos ofrece.
- ~ Cogemos el atributo “humedad”, ya que es el que mayor ganancia de información nos ofrece.

}

Explicación: Para escoger un nodo del árbol a construir, hay que calcular la ganancia de información que nos proporciona cada atributo. Una vez calculadas, Escogemos el atributo que tenga mayor ganancia de información:

Las fórmulas para calcular la Entropía, la Entropía Condicionada y la Ganancia de Información las encontramos en las transparencias 5, 6 y 7, respectivamente.

$$E(\text{Jugar}) = -(9/14) \cdot \log_2(9/14) - (5/14) \cdot \log_2(5/14) = 0.94$$

$$E(\text{Jugar} | \text{Cielo}) = \text{Prob(Nublado)} \cdot E(\text{Jugar} | \text{Nublado}) + \text{Prob(Soleado)} \cdot E(\text{Jugar} | \text{Soleado}) + \text{Prob(Lluvia)} \cdot E(\text{Jugar} | \text{Lluvia})$$

$$\begin{aligned} E(\text{Jugar} | \text{Cielo}) &= (4/14 * 0) + (5/14 * (- (3/5) * \log_2(3/5) - (2/5) * \log_2(3/5))) \\ &+ (5/14 * (- (3/5) * \log_2(3/5) - (2/5) * \log_2(3/5))) = 0.694 \end{aligned}$$

$$\mathbf{IG (\text{Jugar} | \text{Cielo}) = E(\text{Jugar}) - E(\text{Jugar} | \text{Cielo}) = 0.94 - 0.694 = 0.246}$$

$$E(\text{Jugar} | \text{Humedad}) = \text{Prob(Alta)} \cdot E(\text{Jugar} | \text{Alta}) + \text{Prob(Normal)} \cdot E(\text{Jugar} | \text{Normal})$$

Normal)

$$E(\text{Jugar} \mid \text{Humedad}) = (7/14 * (- (4/7) * \log_2(4/7) - (3/4) * \log_2(3/4))) + (7/14 * (- (1/7) * \log_2(1/7) - (6/7) * \log_2(6/7))) = 0.789$$

$$IG(\text{Jugar} \mid \text{Humedad}) = E(\text{Jugar}) - E(\text{Jugar} \mid \text{Humedad}) = 0.94 - 0.79 = 0.151$$

POR LO TANTO EL PRIMER NODO DE NUESTRO ÁRBOL SERÍA EL ATRIBUTO "CIELO"

**Si tenemos la mínima incertidumbre de que el equipo A ganará al equipo B porque el partido está amañado podemos afirmar que{**

- ~ Podemos apostar por A, pero no sabemos nada con certeza
  - = Tenemos máxima información sobre ese partido y podremos apostar seguros por A
  - ~ Deberemos apostar por B; porque seguro que ganara
- }

Explicación: Tener la mínima incertidumbre (0) significa que tenemos la máxima información sobre ese caso; y no se puede dar el caso contrario bajo ningún motivo. Por eso es seguro apostar que el equipo A ganará.

**¿Que queremos conseguir con un árbol de decisión?**

- {
- ~Explorar las infinitas soluciones de un grafo y dar una aleatoriedad.
  - ~Hacer un grafo de 20 nodos.
  - =Saber decidir si tengo que realizar una acción o no.
- }

Explicación: En un árbol de decisión se seleccionan sólo aquellos atributos que nos ofrecen mejor ganancia de información, no todos. Los arboles de decisión nos sirven para determinar si una acción/operación se debe realizar o no.

**¿A qué se debe que el cálculo de la entropía pueda superar el valor 1?:{**

- = A que existen más de 2 valores en la distribución de probables.
  - ~ A que existen 2 o menos valores en la distribución de probables.
  - ~ El cálculo de la entropía nunca supera el 1.
- }

*Explicación: Como vimos en clase, en caso de que tengamos solo 2 posibilidades (el ejemplo "Cara/Cruz de la moneda) tendremos una entropía acotada entre 0 y 1, mientras que a la hora de calcular con más de dos (Ganar/Perder/Empatar...) esta puede rebasar el valor de la unidad.*

**¿Cuál de las siguientes afirmaciones es correcta?:{**

- ~ En los atributos numéricos el ID3 trabaja con valores discretos y continuos.
  - = En los atributos numéricos el ID3 trabaja sólo con valores discretos.
  - ~ En los atributos numéricos el ID3 trabaja sólo con valores continuos.
- }

*Explicación: Como podemos observar en la transparencia 11 el ID3 sólo trabaja con valores discretos. En el caso de que fuesen continuos habría que descomponerlos en rangos.*

**¿Qué afirmación acerca de la entropía es falsa? {**

- ~ En una distribución uniforme la entropía es máxima.
  - ~ La entropía mide el grado de incertidumbre asociado a una distribución de probabilidad.
  - = En una distribución pico la información es mínima.
- }

*Explicación: La falsa es la última puesto que en una distribución pico la entropía es mínima por lo que la incertidumbre es mínima pero la información es la máxima.*

**Los árboles de decisión establecen en qué orden testar los atributos para conseguir la clasificación del vector de entrada. Ahora, ¿cómo se compone dicho orden? {**

- ~ se eligen primero aquellos atributos que peor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.
  - = se eligen primero aquellos atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.
  - ~ se eligen los atributos sin importar que ganancia de información prometen a efectos de descubrir la clase del vector de entrada.
- } Explicación: El orden en el que se testan los atributos de los vectores de entrada para conseguir su clasificación se compone escogiendo primero los atributos que mejor ganancia de información prometen.

**La evaluación de un árbol de decisión se realiza: {**

- = Siempre de derecha a izquierda.
  - ~ Siempre de izquierda a derecha.
  - ~ Ninguna de las anteriores son correctas.
- }

Explicación: Las variables cuyo valor se conoce antes de tomar la decisión , han de aparecer a la izquierda del nodo en el desarrollo del árbol; las que no se conocen aparecerán a la derecha.

**El Algoritmo ID3 que trabaje con atributos continuos tiene que ...{**

- = ...descomponerlos en rangos.
  - ~ ...homogeneizar los rangos por casualidad.
  - ~ ...penalizar los atributos de valor elevado.
- }

Explicación: "ID3 sólo trabaja con atributos discretos. Si se usan atributos continuos hay que descomponerlos en rangos. " (*Transparencia 11 del Tema 6*)

**En un árbol de decisión es cierto que**

- {
- ~ Siempre se seleccionan todos los atributos
  - ~Siempre se seleccionan n-1 atributos, siendo n el número de atributos
  - =No siempre se seleccionan todos los atributos
- }

Explicación: En un árbol de decisión se seleccionan sólo aquellos atributos que nos ofrecen mejor ganancia de información, no todos o n-1, como se puede ver en los ejemplos del tema y como se comentó en clase.

**¿Cual de las siguientes afirmaciones es cierta?(considerando una entropía de distribución de Y condicionada a X){**

- ~ Una entropía condicionada mayor que  $E(Y)$  indica que el conocimiento de X mejora la información que se dispone sobre Y
  - ~ Una entropía condicionada menor que  $E(Y)$  indica que el conocimiento de X no mejora la información que se dispone sobre Y
  - = Ambas son falsas
- }

Explicación: Transparencia 6:

"Una entropía condicionada menor que  $E(Y)$  indica que el conocimiento de X mejora la información que se dispone sobre Y"

**Dado que N es el número de valores que puede tomar una variable, el valor de su entropía se encuentra acotado por: {**

- =  $[0, \log_2(N)]$
  - ~  $[0, \infty)$
  - ~ No es posible determinarlo/ninguna de las anteriores
- }

Explicación:

En el caso de una distribución pico, la entropía es igual a cero (pag. 5 de las transparencias) y, en el caso de una distribución uniforme, los valores de todas las probabilidades son  $1/N$ .

Por lo tanto la entropía es la sumatoria de  $-1/N \cdot \log_2(1/N)$ , dado que todos los términos son iguales, sustituimos la sumatoria por una multiplicación:

$$N(-1/N \cdot \log_2(1/N)) = -\log_2(1/N) = -(\log_2(1) - \log_2(N)) = -(0 - \log_2(N)) = \log_2(N)$$

NOTA: Tal vez sería mas adecuado colocar como respuesta correcta el intervalo  $[0, -\log_2(1/N)]$ , para que no fuese necesario aplicar las identidades logarítmicas.

### **Respecto a la entropía... {**

- ~ En una distribución uniforme, todos los valores son igualmente probables  $P_i = 1/N$  y por tanto la entropía es mínima, lo cual indica máxima incertidumbre.
- ~ Es una medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el valor de otra Y.
- = Ninguna de las anteriores es correcta.

}

Explicación: La entropía es una medida del grado de incertidumbre asociado a una distribución de probabilidad, en una distribución uniforme, todos los valores son igualmente probables  $P_i = 1/N$  y por tanto la entropía es máxima, lo cual indica máxima incertidumbre, por el contrario, en una distribución pico en la que  $P_i = 1$  y  $P_j = 0$ , para todo  $j \neq i$  la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información.

**Sabiendo que el resultado de los partidos disputados entre el Hércules y el Elche ha sido:**

**Gana Hércules(H): 20**

**Empate (X): 5**

**Gana Elche(E): 5**

**Calcula la entropía de que el Hércules gane al Elche en un partido de futbol. {**

~  $E(H) = 0$

=  $E(H) = 1.25$

~  $E(H) = 0.5$

}

Explicación:  $E(H) = -(2/3)(\log(2/3)/(\log 2)) - 2 * ((1/6)\log(1/6)/(\log 2)) = 1.25$

**La fórmula que nos permite calcular la Ganancia Normalizada es: {**

$$G_N(S, A) = \frac{G(S, A)}{\sum_{v_i \in V(A)} p_{v_i} \log_2 p_{v_i}}$$

=

$$G_N(S, A) = \frac{\sum_{v_i \in V(A)} p_{v_i}}{G(S, A) - \sum_{v_i \in V(A)} p_{v_i} \log_2 p_{v_i}}$$

~

$$G_N(S, A) = \frac{G(S, A)}{p_{v_i} \log_2 p_{v_i} - \sum_{v_i \in V(A)} p_{v_i}}$$

~

}

Explicación: La fórmula para calcular la ganancia normalizada es:

$$G_N(S, A) = \frac{G(S, A)}{\sum_{v_i \in V(A)} p_{v_i} \log_2 p_{v_i}}$$

**Referente a los árboles de decisión.... :{**

- ~ es interesante aprenderlos a partir de un conjunto de matrices
- = es interesante aprenderlos a partir de un conjunto de vectores
- ~ ninguna de las anteriores.

}

Explicación: T6, trasparencia 3. Sirven para la clasificación de vectores de atributos (vector de entrada).

**ID3 sólo trabaja con atributos discretos. Para tratar atributos numéricos o continuos hay que...{**

- ignorarlos
- = descomponerlos en rangos
- descomponerlos como máximo en tres valores discretos

}

### Explicacion:

Si se usan atributos continuos hay que descomponerlos en rangos. Para ello se ordenan los ejemplos según el valor y se toman como puntos límite los puntos medios de aquellos en que se cambie de clase. (T6, pag.11)

### **La complejidad de ID3: {**

- ~ ID3 crece linealmente con el número de ejemplos de entrenamiento.
- ~ ID3 crece exponencialmente con el número de atributos.
- = Ambas son correctas.

}

Explicación: El espacio ocupado por id3 es el ocupado por el árbol de decisión. En el peor de los casos, habrá un nodo hoja por cada ejemplo, con un número de nodos intermedios igual al tamaño del espacio de instancias. Esto se produciría por ejemplo, cuando se le pasara como entrada todos los posibles ejemplos, cada uno con una clase diferente. Como intentaría encontrar una forma de separar todos los ejemplos, esto le llevaría a considerar todas las posibles combinaciones de valores de atributos (tamaño de espacio de instancias). Evidentemente, esto no es un caso real. En el mejor de los casos (por ejemplo, todos los ejemplos pertenecen a la misma clase), habrá un único nodo.

### **Se lanza una moneda al aire para ver si sale cara o cruz (dos estados con probabilidad 0,5). Su entropía es: {**

- ~ 0,5
- ~ 0
- = 1

}

Explicación: aplicando la fórmula de la entropía,  $E = 0,5 \cdot \log_2(1/0,5) + 0,5 \cdot \log_2(1/0,5) = (0,5+0,5) \cdot \log_2 2 = 1$

### **Calcular el valor de Ent ([5,2,1]) {**

- = 1.3
- ~ 1.4
- ~ 1.5

}Explicación:  $\text{Ent}([5,2,1]) = -5/8 \log_2 (5/8) - 1/4 \log_2 (1/4) - 1/8 \log_2 (1/8)$

**Según la definición de Ganancia de Información . ¿Cuál de estas respuestas es correcta? :**

- ~ Ganancia de información es conocer el valor de una variable aleatoria X.
- ~ Si obtenemos una alta ganancia implica que el atributo X permite ampliar la incertidumbre de la clasificación.
- = Una alta ganancia implica que el atributo X permite reducir la incertidumbre de la clasificación del ejemplo de entrada. }

Explicación : Ganancia de información es la medida de cuánto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y. Una alta ganancia implica que el atributo X permite reducir la incertidumbre de la clasificación del ejemplo de entrada.

**En los árboles de decisión:** {

- = Se usan algoritmos voraces.
- ~ Vuelta atrás, (Backtracking).
- ~ Ambas son correctas. }

Explicación: No permite la vuelta atrás en los atributos seleccionados. Utiliza el esquema voraz, elegir uno y filtrar recursivamente.

**Si en un dado de 4 caras en el que todas las cara tiene la misma posibilidad de aparecer eliminamos dos caras(nunca pueden salir) que tipo de incertidumbre existiría?**

- {
- =máxima.
- ~mínima.
- ~ninguna de las anteriores.

Explicación: Aunque se eliminen valores de la probabilidad si estos son igual de probable la incertidumbre sigue siendo máxima.

### **Sobre la entropía y el grado de incertidumbre sobre una distribución de probabilidad, ¿Cuál es la opción correcta?:{**

- ~ La entropía es mínima cuando existe una máxima incertidumbre.
  - ~ Cuando una distribución de probabilidad es uniforme la entropía es mínima.
  - = La entropía es máxima cuando existe una máxima incertidumbre.
- }

*Explicación:* La entropía es máxima cuando existe una distribución de probabilidad uniforme ( $P_i = 1/N$ ) o máxima incertidumbre. Es mínima cuando existe una distribución pico ( $P_i = 1$  y  $P_j = 0$ , para todo  $j \neq i$ ) o mínima incertidumbre.

### **En el algoritmo ID3: {**

- ~ Se calcula la ganancia de información de los atributos más prometedores.
- ~ Se escoge el atributo con menor entropía o grado de incertidumbre.
- = *Ninguna de las anteriores.*

}

*Explicación:* se calcula la ganancia de información de todos los atributos y se escoge el de mayor resultado.

### **En una distribución uniforme...:{**

- ~ La entropía es máxima, lo cual indica mínima incertidumbre.
- ~ La entropía es mínima, lo cual indica máxima información.
- = La entropía es máxima, lo cual indica mínima información.

}

*Explicación:* En una distribución uniforme, todos los valores son igualmente probables  $P_i = 1/N$  y por tanto la entropía es máxima, lo cual indica máxima incertidumbre (mínima información).

**¿Cuál de estas afirmaciones es falsa?**

{

~Medida del grado de incertidumbre asociado a una distribución de probabilidad.

=La entropía te da el grado de probabilidad

~La entropía es máxima indica que hay mínima información.

}

Explicación: La entropía te da un grado de incertidumbre, en vez de uno de probabilidad.

**Sobre la entropía en el contexto de IA, se puede decir:{**

= Es la medida del grado de incertidumbre asociado a una distribución de probabilidad.

~ Es interesante aprenderla a partir de un conjunto de vectores.

~ Es siempre una distribución uniforme.

}

*Explicación: La utilización de vectores es normalmente usada para lo estudio de árbol de decisiones. La entropía puede ser una distribución uniforme, pero ni siempre lo es. Luego solamente la primera opción es correcta.*

**Deseamos generar un árbol de decisión para saber si un terreno es apto para viñedo. Para ello partimos de los atributos y valores de la siguiente tabla:**

VARIABLES (Viñedo)					
Casos	Lluvia	Temperatura	Humedad	Fertilidad	Si/No
V <sub>1</sub>	Alta	Irregular	Alta	Normal	No
V <sub>2</sub>	Alta	Irregular	Normal	Alta	No
V <sub>3</sub>	Media	Irregular	Alta	Normal	No
V <sub>4</sub>	Baja	Regular	Alta	Normal	Si
V <sub>5</sub>	Baja	Regular	Alta	Normal	Si
V <sub>6</sub>	Baja	Regular	Normal	Alta	Si
V <sub>7</sub>	Media	Regular	Normal	Alta	Si
V <sub>8</sub>	Alta	Regular	Alta	Normal	No
V <sub>9</sub>	Alta	Regular	Alta	Normal	No
V <sub>10</sub>	Baja	Regular	Normal	Alta	Si
V <sub>11</sub>	Alta	Regular	Alta	Alta	No
V <sub>12</sub>	Media	Irregular	Alta	Alta	Si
V <sub>13</sub>	Media	Irregular	Normal	Normal	No
V <sub>14</sub>	Baja	Irregular	Normal	Normal	No
Totales	Si =	6	No =	8	

**¿Cuál sería el primer atributo del árbol?**

{

~ Temperatura.

~ Fertilidad.

= Lluvia.

}

Explicación: En el siguiente cuadro se pueden observar los cálculos efectuados y vemos que la ganancia del atributo lluvia es mayor que la de cualquier otro.

Entropía Inicial (6+, 8-) =			0,9852			
Valores		Veces	+	-	Entropía	Ganancia
Lluvia	Alta	5	0	5	0,0000	0,4417
	Media	4	2	2	1,0000	
	Baja	5	4	1	0,7219	
	Total/prob.	14	3/7	4/7		
Temperatura	Irregular	6	1	5	0,6500	0,1613
	Regular	8	5	3	0,9544	
	Total/prob.	14	3/7	4/7		
Humedad	Alta	8	3	5	0,9544	0,0113
	Normal	6	3	3	1,0000	
	Total/prob.	14	3/7	4/7		
Fertilidad	Alta	6	4	2	0,9183	0,1281
	Normal	8	2	6	0,8113	
	Total/prob.	14	3/7	4/7		

**El árbol de decisión es utilizado en problemas de... {**

- ~clasificación
- ~inducción
- =ambos son correctos }

Explicación: Los árboles de decisión toman como entrada una situación descrita por un conjunto de atributos y devuelve una decisión que es el predicho valor al valor de entrada.

**Esperamos un mensaje que puede consistir de las letras en minúscula de la a hasta la z.**

**Cual de las siguientes afirmaciones es correcta cuando recibamos el mensaje "qalmnbphijcdgketrsfuvxyzwn̄o"?**

- {
- ~ Es una distribución pico en la que  $P_i = 1$  y  $P_j=0$ , para todo  $j \neq i$  la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información
- = Es una distribución uniforme, todos los valores son igualmente probables  $P_i = 1/N$  y por tanto la entropía es máxima, lo cual indica máxima incertidumbre
- ~ Ninguna de las anteriores es correcta

} Explicación:

Se puede decir que este mensaje llega a nosotros con la máxima entropía (o desorden posible); ya que es poco probable que se pueda pronosticar la entrada de caracteres, pues estos no se repiten ni están ordenados en una forma predecible.

**¿Cuál de las siguientes características de los árboles de decisión es incorrecta? {**

- ~ Es una estructura para la clasificación de vectores de atributos.
- ~ Establece en qué orden testear los atributos para conseguir la clasificación del vector de entrada.
- = Se eligen primero aquellos atributos que menor ganancia de información.

}

Explicación: En los árboles de decisión se eligen primero aquellos atributos con mayor ganancia de información para obtener antes una resolución del problema.

**Atendiendo a los datos de la siguiente tabla:**

X	Y
Moneda	Ca
Azul	ra
Moneda	Ca
Roja	ra
Moneda	Cr
Azul	uz
Moneda	Ca
Verde	ra
Moneda	Cr
Azul	uz
Moneda	Ca
Roja	ra
Moneda	Cr
Roja	uz
Moneda	Cr
Verde	uz

**La Ganancia de Información  $IG(Y|X)$  sería {**

$$= \text{IG}(Y|X) = 0.15625$$

$$\sim |G(Y|X)| = 0$$

$$\sim \text{IG}(Y|X) = 0.84375$$

}

## Explicación:

Al agrupar la información que nos dan obtenemos la siguiente tabla:

$v_j$	Prob ( $X = v_j$ )	$E(Y X = v_j)$
Moneda Azul	3/8	0.875
Moneda Roja	3/8	0.875
Moneda Verde	1/4	0.75

Donde obtenemos las entropías para los tres posibles valores de  $v_i$ :

$$E(Y|X = \text{Moneda Azul}) = -1/4 \cdot \log_2(1/4) - 1/8 \cdot \log_2(1/8) = 0.875$$

$$E(Y|X = \text{Moneda Roja}) = -1/8 \cdot \log_2(1/8) - 1/4 \cdot \log_2(1/4) = 0.875$$

$$E(Y|X = \text{Moneda Verde}) = -1/8 \cdot \log_2(1/8) - 1/8 \cdot \log_2(1/8) = 0.75$$

Por lo tanto para obtener la Entropía Condicionada no hay más que sumar las multiplicaciones de la probabilidad de cada variable por su entropía:

$$E(Y|X) = 3/8 \cdot 0.875 + 3/8 \cdot 0.875 + 2/8 \cdot 0.75 = 0.84375$$

Con lo que ya podemos obtener la Ganancia de Información, previamente calculando la Entropía de Y:

$$E(Y) = -1/2 \cdot \log_2(1/2) - 1/2 \cdot \log_2(1/2) = 1$$

$$IG(Y|X) = 1 - 0.84375 = 0.15625$$

**Se nos plantea resolver un problema de arboles de decisiones mediante el algoritmo ID3, y en la especificación del conjuntos de valores nos encontramos con la siguiente tabla adjunta:**

<b>ID</b>	1	2	3	4	5	6	7	8	9	1
<b>Cantidad</b>	2	1	1	3	5	2	4	4	3	0
	0	5	0	5	0	5	0	5	0	5
	0	0	0	0	0	0	0	0	0	0
<b>Salida</b>	1	1	1	0	1	1	0	1	0	1

Elije la respuesta incorrecta: {

- ~ Habría que ordenar los IDs según la cantidad y tomar como puntos límite los puntos medios de aquellos en que cambie el valor de la salida.
- ~ Estamos ante un caso de atributos numéricos continuos; por lo que es necesario discretizar estos y descomponerlos en rangos.
- = El algoritmo ID3 trabaja con todo tipo de variables, porque el propio algoritmo trata esto sin necesidad de extensiones adicionales.

}

*Explicación: En arboles de decisión trabajan con atributos categóricos y el propio algoritmo ID3 también. En caso de problemas con atributos continuos existen extensiones para este, que tratan precisamente de crear rangos para identificar estos valores y tenerlos así distribuidos de manera categórica.*

**¿Cuál de los siguientes es un algoritmo utilizado en Árboles de Decisión?: {**

- ~ Backtracking
- ~ Ramificación y poda.
- = El algoritmo ID3.

}

*Explicación: El algoritmo ID3 es utilizado en los árboles de decisión, a partir de un conjunto de ejemplos genera unas reglas.*

**¿Cuál de las siguientes características corresponde a un árbol de decisión?:**

{

- = Estructura para clasificación de vectores de atributos.
- ~ Establece en qué orden testar los atributos para conseguir la clasificación del vector de salida.
- ~ Se eligen primero aquellos atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de salida.

}

*Explicación: En las 2 respuestas incorrectas se hace referencia al vector de salida y no al de entrada. Se puede encontrar en la página 3 de las transparencias del tema 6.*

**La entropía puede ser definida como:{**

~Estructura para clasificación de vectores de atributos. ~Medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y. = Medida del grado de incertidumbre asociado a una distribución de probabilidad.

}

*Explicación: la primera opción se refiere a árboles de decisión y la segunda se refiere ganancia de información. La última opción es la correcta.*

**{¿Que es la ganancia de información?:{**

- ~ Estructura para clasificación de vectores de atributos.
  - = Medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y.
  - ~ Medida del grado de incertidumbre asociado a una distribución de probabilidad.
- }

Explicación: La ganancia de información esta definida de esta manera en la transparencia 7 del tema 6

**Respecto a las variables en la clasificación de un arbol de decisión, indica que afirmación es correcta:{**

- ~ Las variables continuas tienen el mismo tratamiento que las variables discretas.
- = Las variables continuas deben establecerse dentro de categorías.
- ~ Ninguna de las anteriores es correcta.

}

Explicación: Las variables continuas deben categorizarse antes de poder ser usadas en un árbol de decisión (deben ser "discretizadas" aunque posteriormente no se seleccionen sus atributos durante el proceso de decisión).

**Indica que afirmación no es correcta:{**

- ~ Ganancia de información es la Medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y.
- ~ Entropía es la medida del grado de incertidumbre asociado a una distribución de probabilidad.
- = Ninguna de las anteriores es correcta.

}

Explicación: Hojas 5 e 7 del Tema 6 contienen las informaciones acerca de GI y Entropía.

**En un sistema dado, en el que todas las opciones tengan la misma probabilidad:** {

- ~ La entropía es baja.
- = La entropía es alta.
- ~ No hay entropía. }

Explicación: Al haber la misma probabilidad para todas las opciones tenemos una total incertidumbre de cual puede ser el resultado. Ejemplo en la transparencia 5 del tema 6.

**En una distribución pico en la que  $P_i = 1$  y  $P_j=0$ ... :** {

=para todo  $j \neq i$  la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información.

~para todo  $j=i$  la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información.

~para todo  $j \neq i$  la entropía es mínima lo cual indica mínima incertidumbre o sea mínima información.

}

*Explicación: En una distribución uniforme, todos los valores son igualmente probables  $P_i = 1/N$  y por tanto la entropía es máxima, lo cual indica máxima incertidumbre. Por el contrario, en una distribución pico en la que  $P_i = 1$  y  $P_j=0$ , para todo  $j \neq i$  la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información.*

**Según la definición de Entropía cuando obtendremos máxima información:**{

- =Cuando hay mínima incertidumbre
- ~Cuando la entropía es máxima
- ~Nunca podemos obtener máxima información

}

Explicación: En una distribución pico en la que  $P_i = 1$  y  $P_j=0$ , para todo  $j \neq i$  la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información.

### **En cuanto a los árboles de decisión:{**

- ~ Es una estructura donde se clasifican vectores de atributos. Estos atributos se testan aleatoriamente.
- ~ Se clasifican vectores de atributos testados en orden. Para componer dicho orden se eligen primero los atributos menos probabilidad tienen de llegar a una solución optima.
- = Se clasifican vectores de atributos testados en orden. Para componer dicho orden se eligen primero los atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.

}

Explicación: Las características de los árboles de decisión son:

Estructura para clasificación de vectores de atributos.

Establece en qué orden testar los atributos para conseguir la clasificación del vector de entrada.

Para componer dicho orden se eligen primero aquellos atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.

Es interesante aprenderlos a partir de un conjunto de vectores

### **Los árboles de decisión son: {**

- ~ Datos sin ningún tipo de relación entre si.
- ~ Una estructura para la separación de vectores por tipos de datos.
- = Una estructura para clasificación de vectores de atributos.

}

*Explicación: Según se muestra en la diapositiva 3 del tema, entre las características de los árboles de decisión, se observa que es una estructura para clasificación de vectores de atributos.*

**La entropía es máxima cuando:** {

- ~ La incertidumbre es mínima.
  - ~ La ganancia de información es máxima.
  - = En una distribución los valores son igualmente probables  $P_i = 1/N$  y por tanto la entropía es máxima.
- }

*Explicación: La entropía es una medida de incertidumbre, lo cual es lo contrario a tener información. Es máxima cuando en una distribución uniforme todos los valores son igualmente probables.  $E(s) = \text{Sumatorio } (-p(i) \log_2(p(i))) = 1$ .*

*Se me ha pasado por alto meter una pregunta en el foro en las fechas adecuadas. No ha sido a propósito, pero ha sucedido. Así que me gustaría que me indicase si directamente estoy suspendido. Un saludo y muchas gracias.*

**Una entropía condicionada menor que  $E(Y)$ :** {

- = Indica que el conocimiento de X mejora la información que se dispone sobre Y.
- ~ Indica que el conocimiento de X reduce la información que se dispone sobre Y.
- ~ No indica nada.

}

**Después de realizar el cálculo de la Entropía condicionada obtenemos que  $E(Y|X) = 0$ . ¿Qué podemos determinar?** {

- ~ Nos hemos equivocado en el cálculo.
- = El conocimiento de X implica el conocimiento de Y.
- ~ El conocimiento de X no nos aporta información alguna sobre Y.

}

*Explicación: Un valor de  $E(Y|X)$  menor que  $E(Y)$  mejora la información que se dispone de Y, pero si  $E(Y|X) = 0$  se puede determinar el valor que tendrá Y dado el valor de X.*

**Tenemos el vector [66,114,66,69,66], ¿cuál es su entropía?**{

= 1,37

~ -1,37

~ 0.486

}

Explicación: Aplicando la fórmula de la entropía (pàg 5 T6) tenemos que  $H = [-3/5 * \log(3/5) - 2/5 * \log(1/5)]/\log2 = 1'37$ . En cualquier caso la segunda respuesta no podría ser por ser negativa y la última sale de cambiar un signo de la fórmula.

**Hallar la ganancia de información con los datos siguientes**

X	Y
P1	Yes
P2	Yes
P3	Yes
P2	Yes
P1	Yes
P2	No
P1	Yes
P3	No
P3	Yes
P2	No

:{

=  $IG(Y|X)=1,243$

~  $IG(Y|X)=0,104$

~  $IG(Y|X)=0,578$

}

Explicación:

**1º Paso:** Hay que conocer cuantos “Yes” y cuantos “No” hay por cada Pm.

P1 → 3 (Yes) y 0 (No)

P2 → 2(Yes) y 2 (No)

P3 → 2 (Yes) y 1 (No)

**2º Paso:** Probabilidad y entropía por cada Pm.

Pm	Prob(X=Pm)	E(Y X = Pm)
P1	0,3	0
P2	0,4	1
P3	0,3	0,918

Cálculos de la entropía por cada Pm:

$$P1: -3/3 \cdot \log_2(3/3) - 0/3 \cdot \log_2(0/3) = 0$$

$$P2: -2/4 \cdot \log_2(2/4) - 2/4 \cdot \log_2(2/4) = 1$$

$$P3: -2/3 \cdot \log_2(2/3) - 1/3 \cdot \log_2(1/3) = 0,918$$

**3º Paso:** Entropía condicionada.

$$E(X|Y) = 0,3 \cdot 0 + 0,4 \cdot 1 + 0,3 \cdot 0,918 = 0,675$$

**4º Paso:** Suma de las entropías de cada Pm.

$$E(Y) = 0 + 1 + 0,918 = 1,918$$

**Último paso:** La ganancia de información.

$$IG(Y|X) = 1,918 - 0,675 = 1,243$$

**La ganancia de información se puede calcular de la siguiente manera:{**

$$= IG(Y | X) = E(Y) - E(Y | X)$$

$$\sim IG(X) = E(Y) - E(Y | X)$$

$$\sim IG(Y | X) = E(X) - E(Y | X)$$

}

*Explicación: La fórmula correcta se puede encontrar en la transparencia 7*

**Si después de lanzar un dado cargado 100 veces tenemos que cada una de las caras ha salido el siguiente número de veces:**

**1 -> 8**

**2 -> 11**

**3 -> 6**

**4 -> 7**

**5 -> 9**

**6 -> 59**

**¿Qué entropía tenemos sobre el lanzamiento de dicho dado? {**

**~ 2.5849**

**= 1.9156**

**~ 1.0152**

**}**

**Explicación: Calculamos la entropía:**

$$E(\text{dado}) = - \frac{8}{100} \log_2(8/100) - \frac{11}{100} \log_2(11/100) - \frac{6}{100} \log_2(6/100) - \frac{7}{100} \log_2(7/100) -$$

$$\frac{9}{100} \log_2(9/100) - \frac{59}{100} \log_2(59/100) = 1.9156$$

Como curiosidad añadir que el primer valor (2.5849) es la entropía que producirá un dado sin cargar, en el que cada valor aparecerá 1 vez de cada 6 lanzamientos, de media.

$$a) E(S) = \sum_{i \in C}^n -p_i \log_2 p_i$$

$$b) E(S) = \sum_{i \in C}^n -p_i + p_i$$

$$c) E(S) = \sum_{i \in C}^n -(p_i \log_2 p_j)^2$$

Teniendo en cuenta las formulas anteriores cual se corresponde al calculo de la Entropía {

= Opción A.

~ Opción B.

~ Opción C.

}

*Explicación: Si observamos en la transparencia 6 del Tema 6, observamos cual es la formula correcta que corresponde con la opción 1*

**¿Qué es la Ganancia de información? :**{

= Medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y.

~ Medida de cuanto ayuda a no conocer el valor de una variable Y para conocer el verdadero valor de otra Y.

~ Ninguna de las anteriores.}

*Explicación: La correcta es la primera opción como podemos ver en la pagina 7 del tema 6*

## TEMA 7 REDES BAYESIANAS

**1. Una red bayesiana permite:**

= Especificar la distribución conjunta de un grupo de variables aleatorias.

**2. La probabilidad tiene...:**

= 3 axiomas

**3. ¿Cuál de las siguientes es un tipo de inferencia en las redes Bayesianas?**

= Aproximada

**4. Indica cuál de las siguientes opciones corresponden con tipos de inferencia en redes Bayesianas:**

= Exacta (Caso general), Casos especiales (Kim&Pearl...), Aproximada..

**5. En las redes bayesianas, ¿para qué queremos la distribución conjunta?**

= A partir de la distribución conjunta podemos contestar cualquier pregunta relativa a la red.

**6. Los mensajes utilizados en el modelo de Kim y Pearl sirven para:**

= Actualizar la credibilidad e introducir nuevas evidencias.

**7. En la inferencia aproximada, los algoritmos de muestreo aleatorio son:**

= Muestreo directo, muestreo por rechazo, MCCM.

**8. Cuál de las siguientes afirmaciones es INCORRECTA acerca de la inferencia exacta general:**

= Uno de los algoritmos que la aplica es el de muestreo directo.

**9. ¿Cuál de las siguientes respuestas es característica el modelo Kim y Pearl?:**

= No existe más de un camino entre cada pareja de nodos

**10. El modelo de Kim y Pearl es un método de inferencia para:**

= Redes bayesianas.

**11. Elige la respuesta correcta:**

= La probabilidad P de un evento a P(a) se define por la frecuencia de a basada en las observaciones pasadas.

El 80% de los jóvenes ya tiene móvil.

a = 'Elegir al azar un joven y que tenga móvil'  $P(a) = 0.8$

**12. Sabemos que:  $P(A|B) P(B) = P(A,B)$   $P(B|A) P(A) = P(B,A) = P(A,B)$**

**La regla de Bayes es:**

$$= \frac{P(A|B) P(B)}{P(B)}$$

**13. La aproximación Bayesiana:**

= Trata de razonar sobre creencias en condiciones de incertidumbre.

**14. Indica cual es la incorrecta sobre el modelo de Kim y Pearl:**

= El método de inferencia son para redes de árboles.

**15. En el ámbito de Redes Bayesianas, decimos que sobre la inferencia exacta en la inferencia aproximada (señala la respuesta correcta) decimos que:**

= Las redes con conexión múltiple son intratables utilizando inferencia exacta.

**16. En probabilidad condicionada,  $P(A|B) = P(A)$  es un suceso:**

= independiente.

**17. Una red bayesiana es:**

= Un grafo acíclico dirigido.

**18. Indica la opción correcta:**

= El modelo de Kim y Pearl es un método de inferencia para redes bayesianas, solo aplicable a un poliárbol.

**19. Si tenemos una variable A = 'Ganador de la liga en el 2022' donde A**

= {a1, a2, a3,....}.

¿Es correcto que  $P(a_1 + a_2 + a_3 + \dots) = 1$ ?

= Si.

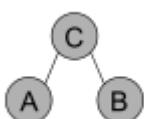
**20. Di cuales NO son tipos de inferencia en redes Bayesianas:**

= Gausiana

**21. Si B representa las variables buscadas, C las conocidas y D las desconocidas, ¿cuál de las siguientes fórmulas corresponde a la regla de inferencia general?**

$$= P(B|C) = a \cdot \sum_D P(B, D, C)$$

**22. Indica que afirmación es correcta en referencia a la siguiente imagen:**



= Si C está relacionada con A y B, y sabemos que se cumplen A y B, al calcular  $P(C|A+, B+)$  la probabilidad de C aumentará.

**23. En la inferencia exacta en poliárboles, el modelo de Kim y Pearl es:**

= únicamente aplicable a un poliárbol

**24. En la probabilidad condicionada, es FALSO que...**

= Si  $P(A|B) = P(A,B)/P(A)$  se dice que son sucesos dependientes.

**25. Dentro de la probabilidad, ¿puede existir la consistencia interna?**

= Sí, es la fiabilidad en sentido estricto.

**26. En la Regla de Bayes, el valor de  $P(B)$  puede ser omitido puesto que en la mayoría de casos:**

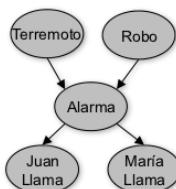
= es un valor constante.

**27. Para responder cualquier pregunta dentro de la red bayesiana empleando el algoritmo de Muestreo Directo visto en clase:**

= Debemos contar las apariciones en  $s[]$  de las evidencias y después debemos dividir por suficientes Muestras

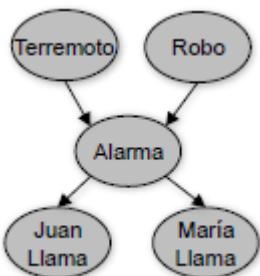
**28. Dada la siguiente imagen que distribución representa:**

$$P(T,R,A,J,M) =$$



$$= P(T) * P(R) * P(A|T,R) * P(J|A) * P(M|A)$$

**29. Dado un suceso  $P(T,R,A,J,M)$  con una distribución de probabilidad conjunta de  $P(T)*P(R)*P(A|T,R)*P(J|A)*P(M|A)$ , podemos afirmar que con una independencia condicional sería:**



$$= 2+2+2^3+2^2+2^2 = 20$$

**30. Selecciona la definición correcta de una Red Bayesiana:**

- = Un grafo acíclico dirigido para representar dependencias entre variables y mostrar una descripción escueta de cualquier distribución de probabilidad conjunta completa.

**31. ¿Qué es la inferencia aproximada?**

- = Son métodos estocásticos, basados en muestreos que simulan las distribuciones de probabilidad de la red.

**32. Segundo como están formadas las redes bayesianas:**

- ~ Cada nodo X tendrá adjunta una distribución  $P(X|Padres(X))$ .
- ~ Si X se conecta con Y se dice que X influencia a Y.
- = Ambas son correctas.

**33. Dada una red bayesiana compuesta por dos variables A y B cuyas probabilidades son  $p(A) = 1/2$ ,  $p(B) = 1/3$  y  $p(A \cap B) = 1/4$ . ¿Cuál sería el resultado de  $p(A | B)$ ?:**

$$= p(A | B) = \frac{1}{3}$$

**34. El algoritmo de Montecarlo resuelve un cálculo sobre una red bayesiana mediante...**

- = ... inferencia aproximada

**35. Si tenemos un problema con conexión múltiple, para tratarlo utilizamos:**

- = Inferencia aproximada.

**36. Dada la siguiente lista de sucesos:  $\{a_1, a_2, a_3, a_4, a_5, a_6\}$ , ¿qué se puede determinar de la siguiente expresión?**

$$\sum_{i=1}^n P(a_i) = 1$$

- = Es cierta

**37. ¿Cuál de las siguientes afirmaciones es la correcta?**

- El modelo de Kim y Pearl sólo se puede aplicar en monoárboles
- Las redes de una sola conexión son intratables utilizando inferencia exacta
- = Las dos anteriores son incorrectas

**38. Dado el siguiente ejemplo:**

Se quiere saber si la familia de un individuo X está en casa basándose en la siguiente información:

- 1.- Si no hay nadie en casa, el perro está fuera.
- 2.- Si el perro tiene problemas de estómago, también permanece fuera.
- 3.- Si el perro está fuera, X oye sus ladridos.
- 4.- El individuo X podría oír ladridos y pensar que son de su perro aunque no fuera así.
- 5.- Si la mujer de X sale de casa, usualmente (pero no siempre) enciende la luz de la entrada.
- 6.- Hay otras ocasiones en las que también enciende la luz de entrada.

Las variables aleatorias (booleanas) en este problema son:

Fuera (nadie en casa), Luz(luz de entrada), Oír(X oye al perro ladrar), Perro(perro fuera), Inst(problemas de estómago del perro)

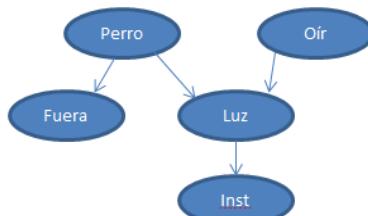
Grafos resultantes:



Grafo A:



Graf.B



Grafo C:

¿Cuál de los siguientes grafos se corresponde con el problema?:

= A

**39. Sean A y B dos sucesos de un espacio de probabilidad tales que:**

**P(A) = 3/5; P(B) = 3/10; P(A∩B) = 1/5; Podemos deducir que:**

=  $P(A|B) = 2/3$

**40. Sabemos que una red bayesiana es un grafo donde los nodos y las aristas que lo componen son, respectivamente:**

- ~ las variables externas a nuestro sistema y dirigidas, representando estas últimas las relaciones existentes entre los nodos.
- ~ las variables de nuestro sistema y no dirigidas, representando estas últimas las relaciones existentes entre los nodos.
- = Ninguna de las anteriores es correcta.

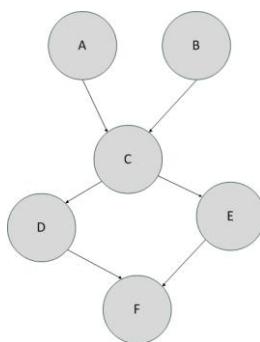
**41. Si disponemos de una moneda trucada, pero no sabemos que lado (cara o cruz) tiene más probabilidad de salir...**

**A: tirada anterior**

**B: tirada actual**

= Si sabemos el valor de la tirada anterior reduciremos nuestra incertidumbre gracias a un conocimiento previo.

**42. ¿Cuál es la distribución conjunta correcta de esta red bayesiana?**



~  $P(A,B,C,D,E,F) = P(A)*P(B)*P(C)*P(D|C)*P(E|C)*P(F|D,E)$

=  $P(A,B,C,D,E,F) = P(A)*P(B)*P(C|A,B)*P(D|C)*P(E|C)*P(F|D,E)$

~  $P(A,B,C,D,E,F) = P(A)*P(B)*P(C|A,B)*P(D|C)*P(E|C)*P(F|D)$

**43. ¿Qué afirmación acerca de las redes bayesianas es falsa?**

- = Están formadas por conjuntos de variables aleatorias que forman nodos independientes entre sí.

**44. El método Montecarlo se basa en la realización de:**

- = Muestreo de cada probabilidad en las sucesivas iteraciones almacenando cada resultado para obtener finalmente una solución basada en aproximaciones

**45. Respecto a las Redes Bayesianas...**

- ~ Su finalidad principal es calcular la distribución conjunta de variables nodo
- ~ Son grafos acíclicos dirigidos para representar dependencias entre variables y mostrar una descripción escueta de cualquier distribución de probabilidad conjunta completa
- = Ambas son correctas

**46. Dada una red bayesiana, su distribución de probabilidad puede expresarse como:**

- ~  $P(x_1, \dots, x_n) = \prod_i P(x_i | \text{pa}(x_i))$ .
- ~  $P(x_1, \dots, x_n) = \prod_i P(x_i | x_{i+1}, \dots, x_n)$
- = Ambas son correctas

**47. En una probabilidad marginal  $P(a) = \sum P(a, b_i)$ , puedo asegurar que:**

- =  $P(a|b) = P(a)$

**48. Sobre el Teorema de Bayes podemos afirmar que:**

- = Nos permite obtener la  $P(A|B)$  en términos de la  $P(B|A)$  siendo  $P(B) > 0$

**49. El problema de la regla de inferencia exacta general es:**

- ~ No funciona para todas las RR.BB.
- ~ Poca complejidad.
- = Ninguna de las anteriores es correcta.

**50. ¿Cuál de las siguientes respuestas es falsa sobre las Redes Bayesianas?:**

- = Su finalidad secundaria es calcular la distribución conjunta de las variables nodo.

**51. Redes con conexión múltiple**

- = son intratables utilizando inferencia exacta

**52. Si un axioma de la probabilidad es:**

**"Si 'a' es un evento cierto, entonces  $P(a) = 1$ " Podemos asegurar que:**

$$= P(a + \neg a) = 1$$

**53. Si Pepe lanza una moneda y luego Juan, ¿qué afirmación es correcta?:**

- ~ Es más probable que Juan saque cara si Pepe ha sacado cruz.
- ~ Es más probable que Juan saque cruz si Pepe ha sacado cruz.
- = Ninguna de las anteriores.

**54. Existen varios tipos de inferencia en Redes Bayesianas, según esto, selecciona la opción correcta:**

- = El modelo de Kim y Pearl es un caso especial de inferencia exacta, en concreto, se denomina inferencia exacta en poliárboles y es un algoritmo más eficiente para tipos específicos de redes.

**55. En una red bayesiana si bajamos por el grafo hasta un nodo terminal, que distribución adjunta tendrá ese nodo:**

$$= P(\text{nodo terminal} | \text{Padres} (\text{nodo terminal})).$$

**56. Sean A y B dos sucesos tales que  $P(B) > 0$ . A es independiente de B cuando:**

$$= P(A|B) = P(A).$$

**57. El modelo de Kim y Pearl:**

- ~ Es solo aplicable a un poliárbol.
- ~ Es un método de inferencia para redes bayesianas.
- = Las dos son correctas.

**58.La principal ventaja de razonamiento probabilística acerca de razonamiento lógico es que...**

= los agentes pueden tomar decisiones racionales mismo cuando no si tiene información suficiente para demostrar que una acción funcionará

**59.Cuál es la finalidad de principal de las redes Bayesianas?**

= Calcular la distribución conjunta de las variables nodo.

**60.De acuerdo con la regla de Bayes,**

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

**es igual a:**

$$= \alpha \cdot P(B|A)P(A)$$

**61.Cuál de las siguientes es incorrecta?**

= Si a y b son mutuamente exclusivos entonces  $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$ .

**62.Cuál de las siguientes afirmaciones acerca de una red bayesiana NO es correcta?**

= Un nodo de una red bayesiana recibe una serie de entradas a través de interconexiones y emite una salida. Esta salida viene dada por tres funciones: función de propagación, función de activación y función de transferencia.

**63.En el modelo de Kim y Pearl para la inferencia en poliárboles existen varios tipos de mensajes, uno de ellos se usa para:**

= Introducir una evidencia.

**64.Según la regla de Bayes ,¿ a que equivale la expresión anterior?:**

$$= P(A | B)$$

**65.Dado R: “Comprar una lavadora” = {r1, r2 ,r3,...} cuál de estas afirmaciones es cierta**

= La probabilidad total  $\sum_{z=1}^n P(r_z) = 1$

**66.** Partimos de que la probabilidad de que llueva un día en concreto es de 0.5 y de que truene es de 0.3. Sabemos además que la probabilidad de que llueva una vez se han escuchado truenos es de 0.2. La probabilidad pues, de que truene una vez que ha empezado a llover es de...

$$= 0.12$$

Explicación: Bayes.  $P(A|B) = (P(B|A) * P(A)) / P(B) = (0.2 * 0.3) / 0.5 = 0.12$

**67.** Teniendo en cuenta la figura de una red bayesiana, indica que respuesta es correcta:

= La probabilidad de que llama María puede ser mayor que 1

**68.** Dado  $P(F,R,A, J, M, C) = P(F) * P(R) * P(A|F,R) * P(J|A) * P(M|A) * P(C|A)$  siendo

Falta de Productos Almacén -> -> Jose

Alarma -> Maria

Robo-----> -

> Carlos

sin independencia condicional{

$$= 2^6 = 64$$

**69.** Una red bayesiana está formada por:

= Un conjunto de enlaces que determinan la influencia (dependencia) entre nodos. Si X se conecta con Y se dice que X influencia a Y

**70.** Dos cajas B1 y B2 contiene 100 y 200 lámparas respectivamente.

La primera caja (B1) tiene 15 lámparas defectuosas y el segundo,

5. Supongamos que una caja es seleccionada al azar y se quita una lámpara. ¿Cuál es la probabilidad de que sea defectuoso?

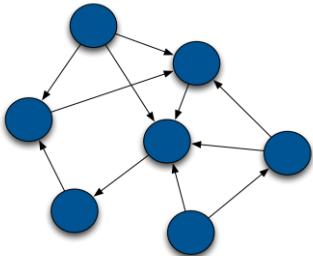
Acerca de:

$$= 9\%$$

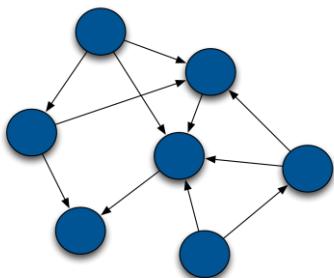
**71.** En cuanto a las redes Bayesianas:

= Está formada por un conjunto de variables aleatorias que forman los nodos de la red. Cada nodo  $x$  tendrá una distribución  $P(X|Padres(X))$

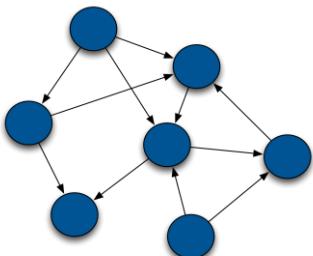
72. De los siguientes grafos, ¿cuál se podría considerar una red bayesiana?



~



=



~

}

73. Supongamos un grupo de personas de las que el 1 % sufre una cierta enfermedad, y el resto está bien. Escogiendo un individuo al azar:

$$P(enfermo) = 1\% = 0.01 \text{ y } P(sano) = 99\% = 0.99$$

Supongamos que aplicando una prueba a una persona que no tiene la enfermedad, hay una posibilidad del 1 % de conseguir un falso positivo, esto es:

$$P(positivo|sano) = 1\% \text{ y } P(negativo|sano) = 99\%$$

Finalmente, supongamos que aplicando la prueba a una persona que tiene la enfermedad, hay una posibilidad del 1 % de un falso negativo, esto es:

$P(\text{negativo}|\text{enfermo}) = 1\%$  y  $P(\text{positivo}|\text{enfermo}) = 99\%$

**¿Cuál es la probabilidad de que un individuo realmente tenga la enfermedad, dado un resultado de la prueba positivo?**

= 50%

**74. ¿Qué desventajas o problemas NO tiene la inferencia exacta?**

= No es posible contestar cualquier pregunta relativa a la red

**75. En un sistema de sucesos independientes si  $P(A|B)=0,5$  Cuál sera la posibilidad de que ocurra el suceso B, teniendo en cuenta que  $P(A,B)$  también es 0,5:**

=  $P(B)$  es 0,5 porque es un suceso independiente

**76. En las redes Bayesianas, el modelo de Kim y Pearl :**

= Solo nos ofrece una inferencia exacta en poliarboles.

**77. Cuál es la finalidad principal de una red bayesiana:**

= Calcular la distribución conjunta de las variables nodo

**78. Cuál de estas afirmaciones en relación al modelo de Kim y Pearl es INCORRECTA:**

= solo es aplicable para políárboles y monoárboles.

**79. ¿Qué diferencia existe entre la aproximación frecuencial y la aproximación bayesiana?**

~ La aproximación frecuencial utiliza los eventos pasados para predecir los presentes.

~ La aproximación bayesiana razona sobre creencias en condiciones de incertidumbre.

= Ambas son correctas.

**80. Una red bayesiana es:**

= Un grafo acíclico dirigido para representar dependencias entre variables y mostrar una descripción escueta de cualquier distribución de probabilidad conjunta completa.

**81. Las redes bayesianas están formadas por:**

- ~ Un conjunto de variables aleatorias que forman los nodos de la red. Cada nodo X tendrá adjunta una distribución  $P(X|Padres(X))$ .
- ~ Un conjunto de enlaces que determinan la influencia (dependencia) entre nodos. Si X se conecta con Y se dice que X influye a Y.
- = Ambas respuestas son correctas.

**82. ¿Cuál es el problema de la inferencia exacta general?**

- = Mucha complejidad.

**83. El modelo de Kim y Pearl:**

- = Se basa en el paso de dos tipos de mensajes entre nodos.

## **TEMA 8. APRENDIZAJE BAYESIANO**

**1. El teorema de Bayes aplicado al aprendizaje permite:**

= Conocer el máximo a posteriori MAP

**2. Cuántos y qué tipos de aprendizaje podemos encontrarnos en el aprendizaje automático:**

= 3 tipos: supervisado, no supervisado y por refuerzo.

**3. En el aprendizaje bayesiano, el máximo a posteriori o MAP es:**

=  $h_{MAP} = \text{argmax}_h P(h|D)$

**4. ¿Por cuántas fases pasa el aprendizaje bayesiano?:**

= Por 2 fases.

**5. El clasificador bayesiano "naive" asume que:**

= los atributos que describen a los ejemplos son condicionalmente independientes entre sí con respecto al concepto que se pretende aprender

**6. Considerando 3 variables A=1, B=5, C=2, el resultado devuelto por  $\text{argmax}(A, B, C)$  es :**

= B

**7. Respecto a los fundamentos MAP y ML es cierto que :**

= ML implica máxima verosimilitud, debido a que  $P(h) = \text{cte.}$

**8. Respecto al aprendizaje bayesiano, ¿cuál de las siguientes opciones es correcta?:**

= Permite combinar los datos de ejemplo con conocimiento a priori.

**9. En el Aprendizaje Supervisado**

= Hay que entrenar hasta alcanzar el error de validación mínimo.

**10. El conjunto de validación en un proceso de aprendizaje supervisado sirve**

= Para comprobar la capacidad de generalización del proceso y evitar el sobreentrenamiento.

**11. En el aprendizaje supervisado, la regresión:**

= Estima una función continua.

**12. Respecto a las hipótesis en el clasificador bayesiano:**

~ El ejemplo el cual pertenece a la hipótesis está caracterizado como tuplas de atributos.

~ Las hipótesis son las clases a las que puede pertenecer un ejemplo

= Las dos respuestas son correctas

**13. ¿Qué permite hacer el aprendizaje bayesiano?:**

= permite combinar los datos de ejemplo con conocimientos a priori

**14. De los siguientes tipos de aprendizaje, indica la definición correcta:**

=Aprendizaje supervisado: Aprendemos a partir de ejemplos conocidos (etiquetados según su clase).

**15. El decisor de máxima verosimilitud, ML, asume que todas las hipótesis son equiprobables:**

= A priori

**16. El uso de reconocimiento de patrones a partir de ejemplos conocidos es una característica del:**

= Aprendizaje supervisado.

**17. ¿En qué consiste el aprendizaje automático? :**

= Consiste en programar una computadora para que mejore en la realización de una tarea a partir de datos de ejemplo o de la experiencia.

**18. El aprendizaje por refuerzo, cuál de las siguientes es correcta:**

=No parte de un conjunto de datos ejemplo.

**19. En la fase 1 de aprendizaje voc significa:**

= Conjunto de palabras en X (sin repetición y sin considerar preposiciones, artículos, etc...)

**20. Dada esta definición: "Es un método de aprendizaje Automático donde un modelo es ajustado a las observaciones y no tiene ningún conocimiento a priori" a cuál de las siguientes corresponde**

=Aprendizaje no supervisado

**21. El conjunto de validación:**

= Se emplea para estimar el error de generalización.

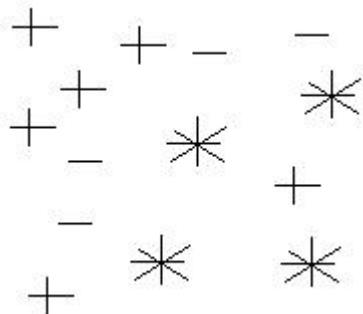
**22. En el aprendizaje supervisado, aleatoriamente se parte el conjunto inicial de ejemplos en:**

= dos grupos ,conjunto de entrenamiento y conjunto de validación.

**23. ¿Cómo solucionamos el problema de sobreentrenar un algoritmo de aprendizaje?**

= A través del error de validación.

**24. Dada la siguiente imagen y teniendo en cuenta que todos los atributos están etiquetados y clasificados, podemos decir que:**



{

= Tiene 3 tipos clases, una dimensión por cada atributo y pertenece al tipo de aprendizaje Supervisado

**25. Un ejemplo de área de aplicación del aprendizaje automático es:**

= Minería de datos (Data mining)

**26. Ante un problema, aplicaremos el clasificador bayesiano naive cuando:**

~ Dispongamos de conjuntos de entrenamiento de tamaño medio o grande.

~ Los atributos que describen a los ejemplos sean independientes entre sí con respecto al concepto que se pretende aprender.

= Ambas son correctas.

**27. ¿Cuál de los siguientes conjuntos del aprendizaje supervisado no es el correcto?:**

= Cluster

**28. Cuando los conjuntos de entrenamiento no son suficientemente grandes utilizamos la técnica de validación cruzada k-fold cross validation, que funciona:**

= Coge un conjunto y lo divide en k conjuntos del mismo tamaño y hace un conjunto para validación y los otros para entrenamiento, cambiándose los papeles para que todos sean validadores y de entrenamiento.

**29. En el clasificador bayesiano:**

- ~ Las hipótesis son las clases a las que puede pertenecer un ejemplo.
- ~ Suponemos ejemplos caracterizados como tuplas de atributos  $\langle a_1, a_2, \dots, a_n \rangle$
- = Las dos son correctas.

**30. Cuál de las siguientes afirmaciones es correcta:**

- ~ En el aprendizaje supervisado, en caso de no haber un gran número de ejemplos, se aplica la técnica k-fold cross validation
- ~ En el aprendizaje supervisado, se divide el conjunto inicial de ejemplos en dos grupos, uno para entrenar y otro para la validación
- = Ambas son correctas

**31. Indica cuál de las siguientes afirmaciones sobre el teorema de bayes es correcta:**

- = Es utilizado para detectar si un correo es spam o no mediante palabras guardadas en un vocabulario.

**32. Respecto al clasificador bayesiano podemos decir que es correcto que:**

- = Las hipótesis son las clases a las que puede pertenecer un ejemplo.

**33. Cuando los conjuntos de entrenamiento no son lo suficientemente grandes usamos la técnica de...:**

- ~ Validación paralela.
- ~ Conjunto de validación.
- = Ninguna de las anteriores es correcta.

**34. Dentro del tipo de aprendizajes, cual es el que agrupa los datos en clusters**

- = Aprendizaje no supervisado

**35. Cuando estamos en la fase de aprendizaje en un clasificador de textos usando naive bayes, por qué aparece el 1 en el numerador y el término  $|Voc|$  en el denominador del cálculo de la probabilidad de una palabra en una categoría dada:  $P(w_k|c_j) = (n_k+1)/(n+|Voc|)$ :**

- = Para evitar que la probabilidad salga 0 si la palabra no ha aparecido nunca.

**36. Estamos implementando un modelo de aprendizaje para guiar a nuestro robot autómata "Emilio" en un entorno laberintico mediante sucesivas pruebas a base de prueba/error; y utilizando simplemente 3 reglas de movimiento, las cuales impiden retroceder en el mapa, y que son: izquierda, adelante y derecha. Únicamente podemos avanzar, de modo que no podemos ir hacia atrás en el mapa, ni usando una regla específica (como se ha comentado), ni usando giros a la izquierda o derecha. Sabiendo esto, indica que esquema de aprendizaje se adaptaría más al modelo planteado:**

- ~ Aprendizaje supervisado
- ~ Aprendizaje NO supervisado
- = Ninguna de las anteriores

**37. El aprendizaje de bayesiano usado como clasificador:**

- = puede obtener probabilidades de pertenecer a cada clase.

**38. ¿Qué necesidades impulsan la idea del aprendizaje automático?:**

- = Ambas son correctas.
- ~ Situaciones o tareas donde el algoritmo que se requiere debe adaptarse a circunstancias particulares.
- ~ Algoritmos difíciles de programar "a mano".

**39. Sea el ejemplo de la predicción del tiempo visto en clase:**

DÍA	TEMP.	DIR. VIENTO	CIELO	PRESIÓN	Tiempo
1	$\leq 0$	Sur	Nuboso	Subiendo	Sol
2	$> 0$	Oeste	Claro	Estable	Sol
3	$> 0$	Norte	Claro	Subiendo	Sol
4	$> 0$	Norte	Claro	Bajando	Lluvia
5	$> 0$	Oeste	Nuboso	Bajando	Lluvia
6	$\leq 0$	Norte	Nuboso	Bajando	Nieve
7	$> 0$	Sur	Nuboso	Estable	Lluvia
8	$> 0$	Sur	Claro	Subiendo	Sol
9	$\leq 0$	Este	Nuboso	Bajando	Nieve
10	$\leq 0$	Sur	Claro	Estable	Sol

Podemos decir que la predicción del tiempo para  
 $\langle \text{presión}=\text{Subiendo}, \text{cielo}=\text{Nuboso} \rangle$   
sería:

= Sol

**40. Un alumno de la UA guarda los resultados de sus notas finales en 1ª convocatoria de las asignaturas que ha cursado, así como si la asignatura era de primer o segundo cuatrimestre, si el profesor que le impartía la teoría era titular o asociado y la base de conocimiento de dicha asignatura (Programación, sistemas o teoría de la información) en la siguiente tabla:**

Asignatura	Base de Conocimiento	Cuatrimestre	Profesor	Nota final
A	Sistemas	Primero	Titular	Aprobado
B	Programación	Primero	Titular	Sobresaliente
C	Sistemas	Primero	Titular	Notable
D	Teoría de la información	Primero	Asociado	Suspensos
E	Programación	Primero	Titular	Sobresaliente
F	Sistemas	segundo	Titular	Aprobado
G	Teoría de la información	segundo	Asociado	Notable
H	Programación	segundo	Asociado	Sobresaliente
I	Programación	segundo	Titular	Suspensos
J	Sistemas	Primero	Titular	Notable

K	Teoria de la información	Primero	Titular	Aprobado
L	Programación	Primero	Asociado	Sobresaliente

**Si para este curso se ha cogido una asignatura de programación de segundo cuatrimestre y la teoria se la imparte un profesor asociado ¿Qué calificación es más probable que obtenga en esta asignatura?**

= Sobresaliente.

**41. ¿Qué tipo de aprendizaje funciona a partir de ejemplos conocidos? :**

= Aprendizaje supervisado

**42. En el aprendizaje bayesiano durante la fase 1 (aprendizaje), después de tomar un conjunto de ejemplos  $xi \in X$  etiquetados con las clases a las que pertenecen, debemos calcular la probabilidad a priori  $P(cj)$  para cada clase  $cj$ , lo cual se realiza de la siguiente manera:**

=  $P(cj) = n^o$  de ejemplos etiquetados con  $cj$  /  $n^o$  total de ejemplos

**43. En el aprendizaje supervisado:**

= conocemos la clase a la que pertenece cada ejemplo.

**44. La minería de datos intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de BD . Consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior.**

**Podemos decir que NO utiliza el tipo de aprendizaje:**

= Refuerzo.

**45. Dados los tipos de aprendizaje y en concreto el Aprendizaje por Refuerzo, señala la afirmación correcta:**

= En el Aprendizaje por Refuerzo tenemos una medida de lo bien o mal que está funcionando el algoritmo, pero no sabemos exactamente qué falla.

**46.**

Día	tráfico	pasajeros	tiempo	llegada del autobus
1	poco	muchos	lluvia	<b>puntual</b>
2	normal	pocos	sol	<b>puntual</b>
3	congestión	muchos	sol	<b>retrasado</b>
4	normal	pocos	nieve	<b>retrasado</b>
5	normal	muchos	lluvia	<b>retrasado</b>
6	poco	pocos	sol	<b>temprano</b>

**Cuáles son las predicciones correctas de la llegada del autobús para <pasajeros = muchos, tiempo = sol> utilizando  $P(c_i) \prod_j P(a_j / c_i)$ ?**

=  $c_i = \text{puntual} \rightarrow 1/12, c_i = \text{temprano} \rightarrow 0, c_i = \text{retrasado} \rightarrow 1/9$

**47. En el Aprendizaje por refuerzo**

= tenemos una medida de lo bien o mal que está funcionando el algoritmo, pero no sabemos exactamente qué falla.

**48. Acerca de las características de las redes de aprendizaje bayesianas, es dicho:**

- ~Un conocimiento a priori puede ser combinado con los datos observados para determinar la probabilidad de una hipótesis.
- ~Los métodos bayesianos pueden acomodar las hipótesis que hacen predicciones probabilísticas.
- = Ambos son correctas.

**49. Dadas las variables X=9, Y=3, Z=5, la solución a  $\max(X, Y, Z)$  y a  $\text{argmax}(X, Y, Z)$  es :**

= 9 y X respectivamente.

**50. Un algoritmo de aprendizaje:**

= es consistente si obtiene una hipótesis que no comete ningún error sobre los ejemplos de entrenamiento.

**51. En el tipo de aprendizaje no supervisado tenemos:**

= un conjunto de datos que queremos agrupar en clusters.

**52. Despues de realizar el cálculo del siguiente caso de predicción meteorológica, obtenemos  $P(T<0, \text{Viento=Norte} | \text{Lluvia})=0$  ¿Qué podemos determinar?**

= No lloverá.

**53. Cuál de las siguientes afirmaciones acerca del aprendizaje bayesiano es falsa:**

= Esta basado en el teorema -->  $P(h|D) = P(D) / P(D|h)P(h)$

**54. Google completa la cadena que introduces en la barra del buscador, utiliza el modelo oculto de markov que es un modelo estadistico cuyo objetivo es determinar los parámetros desconocidos de dicha cadena a partir de los parámetros observables, permite combinar los datos de la cadena con conocimiento a priori.**



uni

universidad de alicante

universidad de murcia

universidad miguel hernandez

**Podemos decir que :**

= Se basa en el tipo de aprendizaje supervisado.

**55. Cuál de las siguientes afirmaciones es correcta, sobre los aprendizajes automáticos:{**

= En el aprendizaje supervisado el error de generalización se estima mediante un conjunto de validación.

**56. Respecto al aprendizaje no supervisado, donde tenemos un conjunto de datos que queremos agrupar en clusters:**

= No tenemos etiquetas de a que pertenecen los datos.

**57. En una competición de tiro con arco después de 10 lanzamientos de un competidor se han registrado los siguientes datos:**

Tiro	Viento	Distancia (m)	Puntuación
1	SUR	50	30
2	ESTE	60	10
3	SUR	50	20
4	NORTE	60	30
5	NORTE	50	10
6	ESTE	75	30
7	SUR	60	30
8	SUR	60	10
9	OESTE	50	30
10	ESTE	60	20

**Calcula la predicción de la puntuación para <viento = SUR, distancia = 60>**

=  $c(30) = 0.08$ ,  $c(20) = 0.05$ ,  $c(10) = 0.0666$

**58. Según el aprendizaje bayesiano ¿Cómo se estima una probabilidad?**

=  $n^o$  de ejemplos de la clase actual /  $n^o$  total de ejemplos

## **59. Acerca de las dificultades del aprendizaje bayesiano**

~ Necesidad de un conocimiento a priori. Si no se tiene este conocimiento estas probabilidades han de ser estimadas

~Coste computacional alto. En el caso general es lineal con el número de hipótesis candidatas.

= Ambas son ciertas

## **60. ¿Cuál es la idea del aprendizaje automático?**

= En programar una computadora para que mejore en la realización de una tarea a partir de datos de ejemplo o de la experiencia.

## **61. En "naive bayes" el cálculo simplificado consiste en:**

= Calcular el productorio de probabilidades de todos los atributos del ejemplo a clasificar condicionado a la clase que se está probando.

## **62. ¿Cuál de estas opciones es correcta con respecto a la simplificación?:**

= Suponemos que los valores de los atributos son condicionalmente independientes para una clase dada

## **63. ¿Cuál de estas opciones es característica del Aprendizaje Bayesiano?:**

= Es necesario conocimiento a priori.

## **64. ¿Cuál de las siguientes afirmaciones no corresponde al aprendizaje bayesiano?:**

= Permite combinar los datos de ejemplo con conocimiento a posteriori

**65. ¿En cuál de las aprendizajes se considera que el problema de la construcción de hipótesis a partir de datos es un subproblema del problema más importante relacionado con la formulación de predicciones?:**

= Aprendizaje bayesiano

**66. Sabemos que el teorema de Bayes aplicado al aprendizaje permite conocer el máximo a posteriori MAP, por lo tanto, podemos señalar como opción incorrecta:**

= El aprendizaje bayesiano permite combinar los datos de ejemplo con conocimiento a posteriori.

**67. El aprendizaje bayesiano se basa en el teorema de Bayes, el cual es:**

=  $P(h|D) = P(D|h) * P(h) / P(D)$ .

**68. Respecto al aprendizaje bayesiano indica la opción correcta:**

- ~ En la fase 3 hallamos la estimación MAP.
- ~ En la fase 2 calculamos  $\text{docs}_j$ , documentos de la clase  $c_j$ .
- = Ninguna de las anteriores.

**69. En qué consisten las simplificaciones:**

= No importa el orden de aparición de los elementos, sólo importa si estos elementos están presentes.

**70. Los datos de un neumático de un Fórmula 1 en una carrera son los siguientes:**

Vuelta	Temperatura	Estado	Presión	Tiempo
1	30	Seco	Media	Soleado
2	30	Seco	Alta	Soleado
3	29	Seco	Alta	Nublado
4	30	Seco	Alta	Soleado
5	30	Seco	Media	Lluvia
6	29	Mojado	Baja	Lluvia
7	29	Mojado	Baja	Lluvia
8	30	Mojado	Baja	Nublado
9	31	Mojado	Baja	Soleado
10	30	Mojado	Media	Soleado
11	31	Mojado	Baja	Soleado
12	32	Seco	Alta	Soleado

**La predicción de una alta presión del neumático cuando el estado es mojado y el tiempo soleado es:**

= 0

**71. En el aprendizaje bayesiano, al estimar las probabilidades:**

= Suponemos que los valores de los atributos son condicionalmente independientes para una clase dada.

**72. Si hablamos del aprendizaje bayesiano, en la fase de clasificación:**

= Dado un documento  $x=w_1, w_2, \dots, w_n$ , nos quedamos con las posiciones de palabras que están contenidas en  $Voc$  y devolvemos la estimación  $map$ .

**73. ¿Cuál de los siguientes afirmaciones acerca de problemas del aprendizaje bayesiano NO es correcta?**

= No hay la posibilidad de construir representaciones más complejas.

**74. En el aprendizaje Bayesiano existe Máxima Verosimilitud (ML) cuando:**

=  $P(h) = cte$

**75. ¿Qué presupone el clasificador bayesiano “naive”?:**

- = Que todos los atributos tienen que ser independientes

**76. La siguiente fórmula es la base del aprendizaje bayesiano:**

$$h_{MAP} = \text{argmax}_h P(h | D)$$

**¿Qué significa?:**

- ~ Buscar la hipótesis h con el valor máximo a posteriori.
- ~ Buscar la hipótesis h más probable si hemos observado una serie de datos D.
- = Las tres respuestas son correctas.

**77. Podemos calcular la probabilidad de los atributos condicionada a las diferentes categorías**  $P(a_1, a_2, \dots, a_n | c_i)$  **con la aproximación**  $\prod_{j=1}^n P(a_j | c_i)$  **considerando:**

- = Los atributos son independientes entre sí.

**78. En la fase 1, Aprendizaje, de Clasificador bayesiano “naive”, porque se añade el 1 a nk (nk+1)**

- = Para que cuando la probabilidad dé 0 normalmente, ese valor 0, no influe demasiado en el resultado final.

**79. ¿Qué método de aprendizaje utilizaríamos para hacer que un robot con un brazo mecánico pudiese mejorar jugando a ping pong?:**

- = Aprendizaje por refuerzo

**80. El aprendizaje bayesiano:**

- = Basado en el teorema de Bayes.

**81. El aprendizaje bayesiano consiste en:**

=Buscar la hipótesis  $h$  más probable si hemos observado una serie de datos  $D$ .  $h_{MAP} = \arg\max P(h|D)$

**82. Es falso que un clasificador Bayesiano:**

= No necesita ningún tipo de conocimiento previo para realizar la clasificación.

**83. ¿Qué tipo de aprendizaje tiene una medida de lo bien o lo mal que está funcionando el algoritmo, pero no sabemos exactamente lo que falla?**

= Aprendizaje por refuerzo.

**84. En cuanto al aprendizaje bayesiano...**

= Permite combinar los datos de ejemplo con reconocimiento a priori.

**85. En un ejemplo clasificador de textos, tenemos que clasificar un texto en una categoría predefinida dado cualquier conjunto de palabras de un texto (atributos), en las simplificaciones podemos afirmar que: {**

= La posición de las palabras en el texto importa o no dependiendo del objetivo.

## TEMA 9 SI

Acerca de las neuronas{

- ~ Son dispositivos de "todo o nada".
- ~ Las salidas que provoca pueden afectar al resultado final.
- = Ambas son ciertas

}

El algoritmo Backpropagation para redes neuronales {

- ~ Se aplica para un grupo muestra del conjunto de entrenamiento
- ~ Calcula la salida de la red en la última fase
- = Itera hasta que el error baje de un umbral.

}

Respecto a las funciones de activación derivables:{

- ~ Para aplicar el algoritmo de entrenamiento multicapa no es necesario que la función de activación sea derivable, aunque es conveniente que lo sea.
- = Buscamos funciones derivables con forma similar al escalón del perceptrón de una sola capa.
- ~ Ambas son correctas.

}

Respecto a la convergencia de backpropagation:{

- = Una red neuronal converge cuando el error de validación se mantiene bajo y los ejemplos de entrenamiento no provocan cambios significativos en los pesos de la red.
- ~ Una red neuronal converge cuando el error de validación se mantiene alto y los ejemplos de entrenamiento no provocan cambios significativos en los pesos de la red.
- ~ Una red neuronal converge cuando el error de validación se mantiene bajo y los ejemplos de entrenamiento provocan cambios significativos en los pesos de la red.

}

La Inicialización de los pesos de la red es la siguiente:

- Arbitraria
- Aleatoria

¿Cual es el problema de la inicialización en los descensos por gradiente?{

= Los mínimos locales.

~ El entrenamiento de la red.

~ No tiene ningún problema .

}

El overshooting ocurre{

~ Cuando nos saltamos el máximo.

~ Cuando la media calculada es mayor que el valor previamente calculado en la iteración anterior.

= Cuando nos saltamos el mínimo

}

Respecto al problema de los mínimos locales en la inicialización en los descensos por gradiente, la solución aportada a dicho problema es: {

~ El problema de mínimos locales se encuentra solo en los ascensos por gradiente, por lo que no requiere solución para los descensos por gradiente.

~ Inicializar los pesos de la red con un valor constante.

= Entrenar la red desde las distintas inicializaciones.

}

En redes neuronales, cuando no hay separación lineal, ¿cómo se resuelven los problemas de paridad? {

~ Para resolver estos problemas es preciso eliminar una capa de neuronas.

= Para resolver estos problemas es preciso incorporar una capa adicional de neuronas.

~ Para resolver estos problemas es preciso incorporar dos capas adicionales de neuronas.

}

En el Entrenamiento, el ajuste de hiperplanos ocurre: {

= Dados dos conjuntos de ejemplos correspondientes a dos clases, buscaremos su separación por un hiperplano.

~ Dados X conjuntos de ejemplos correspondientes a X clases, buscaremos su separación por un hiperplano.

~ Ninguna de las anteriores.

}

Respecto a la Regla delta: {

~ Permite ajustar iterativamente el hiperplano.

~ Se asume que el incremento de los pesos es proporcional a la disparidad entre la salida observada y la salida deseada..

= Todas las respuestas son correctas..

}

¿Qué es necesario para aplicar el algoritmo de entrenamiento multicapa?{

~ Que la función de activación no sea derivable.

= Que la función de activación sea derivable.

~ Es indiferente si se puede derivar o no.

}

Sobre la interpretación geométrica, cuál de las siguientes afirmaciones es falsa: {

= Problemas con regiones de decisión mas complejas no requieren distintas estrategias de separación.

~ Dichas estrategias las proporcionan las capas ocultas

~ Problemas con regiones de decisión mas complejas requieren distintas estrategias de separación.

}

En la convergencia de backpropagation, el ajuste de la constante  $\eta$ {

= Valores muy pequeños: convergencia lenta.

~ Valores muy pequeños: convergencia rápida.

~ Valores muy pequeños: no converge.

}

Si en la neurona biológica las entradas serían las dendritas en la neurona computacional las entradas son:{

~ solo números naturales

= números reales

~ ninguna de las anteriores

}

Acerca de los perceptrones multi-capas, cual de las siguientes es INCORRECTA:

{

= La "interpretación geométrica" se desarrolló sobre un algoritmo que permite encontrar los pesos asociados a todas las neuronas.

~ Un perceptrón con dos capas ocultas es capaz de aproximar cualquier función.

~ La capa adicional se denomina capa oculta.

}

En la inicialización de la red ¿Qué problema se pretende evitar?: {

= Mínimo local.

~ Máximo global.

~ No hay que evitar nada.

}

La inicialización de pesos en una red neuronal artificial debe de realizarse: {

= De forma aleatoria o arbitraria.

~ Igual que el número de neuronas.

~ Debe ser inicializada a 0.5, pues se ha comprobado de forma empírica que se obtienen los mejores resultados.

}

Que afirmación acerca de la redes neurales es falsa {

= El modelo computacional no se asemeja en nada a una neurona biológica.

~ Las entradas de una neurona computacional son números reales.

~ Una neurona biológica consta básicamente de 3 partes: entradas, integración y salidas.

}

Un perceptrón puede aproximar cualquier función: {

~ Siempre que la función sea derivable y que posea una capa oculta

~ Siempre

= Sólo si posee dos capas ocultas

}

**¿Cuáles son las fases del algoritmo *Backpropagation* ?{**

~ Hacia delante y actualización de los pesos de todas las capas.

~ Hacia atrás y actualización de los pesos de todas las capas.

= Hacia delante, hacia atrás y actualización de los pesos de todas las capas.

}

Una neurona computacional consta de: {

= Entrada, integración y salida.

~ Entrada, acoplamiento y salida.

~ Dendritas, soma y axón

}

Si estamos ante una red neuronal artificial con perceptrones multicapa, y en las neuronas no se produce procesamiento; podemos afirmar que nos encontramos en: {

= La capa de entrada.

~ Las capas ocultas.

~ La capa de salida.

}

¿Cuál de las siguientes afirmaciones es correcta?{

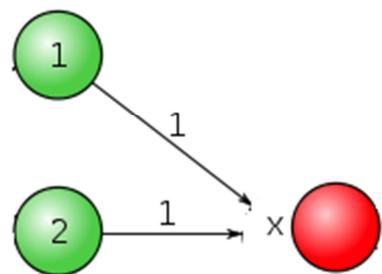
~ Una neurona nos puede decir si va a llover, nevar o sol (elección de tiempo).

= Una neurona nos puede decir si es que va a llover o no (elección de si o no).

~ Ambas respuestas son correctas.

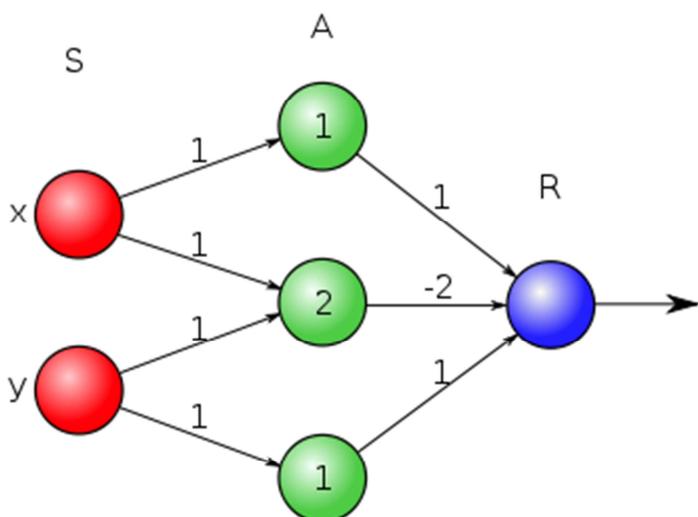
{

La representación de la función booleana AND mediante perceptrones es (Indica la respuesta correcta): {

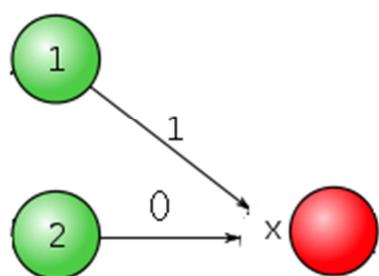


=

~



~



$$\delta_k = (d_k - y_k) * f'(net_k)$$

La formula anterior, ¿qué es lo que mide? {

= Mide lo que contribuye cada neurona al error obtenido.

~ Mide el nuevo estado de activación de la neurona.

~ Mide la mejor salida de la red.

}

Una red neuronal converge cuando: {

= El error de validación se mantiene bajo y los ejemplos de entrenamiento no provocan cambios significativos en los pesos de la red.

- ~ Dicha neurona es responsable de un error.
  - ~ Ninguna de las anteriores.
- }

En la interpretación geométrica de la función de activación, los ejemplos que cumplan la siguiente fórmula:

$$\sum_{i=1}^n x_i w_i - \theta = 0$$

¿Pertenecen al hiperplano? {

- = Sí
  - ~ No
  - ~ Son ejemplos indefinidos
- }

El problema de la XOR es un problema: {

- = de paridad.
  - ~ relacionado con la constante de aprendizaje.
  - ~ de gradiente.
- }

Las redes neuronales artificiales pueden sufrir de sobreentrenamiento y arrojar resultados erroneos:{

- = debido al exceso de flexibilidad o rigidez.
  - ~ las redes neuronales no pueden sobreentrenarse.
  - ~ debido al uso de un número insuficiente de neuronas en la capa oculta.
- }

**En una neurona computacional, la integración es:**

{

= la suma ponderada (net) por los pesos sinápticos seguida de una función de activación  $f(\text{net})$ .

~ la suma ponderada (net) por los pesos sinápticos seguida de una función de desactivación  $f(\neg\text{net})$ .

~ Ambas son correctas.

}

Entradas, integración y salidas son las partes de...{

~ neurona biológica

~ neurona computacional

= ambos tipos de neuronas

}

Teniendo presente que  $\Delta w_i = \eta(d - y)x_i$  siendo **d** la salida deseada e **y** la observada, podemos afirmar que:{

= Con el valor de entrada 1 y valor de salida 0 la red falla.

~ Con el valor de entrada 0 y valor de salida 1 la red acierta.

~ Con el valor de entrada 1 y valor de salida 1 la red falla.

}

En cuanto al error  $\delta$  de las Redes Neuronales {

= Una neurona de una capa intermedia contribuye en los  $\delta$  de las de la capa siguiente.

~ Los  $\delta$  de cada capa son independientes entre si y no afectan al resultado de las siguientes.

~ El cálculo de los  $\delta$  no es necesario en el algoritmo Backpropagation.

}

Para la backpropagation necesitamos:{

= Una función derivable.

- ~ Una función no derivable.
  - ~ Podemos usar cualquier función
- }

De las siguientes afirmaciones cuál es la correcta: {

- ~ La convergencia en las redes neuronales se refiere a que el valor del error de validación se mantiene constante.
  - ~ La integración en una neurona computacional se refiere a la suma ponderada por los pesos sinápticos.
- = Ninguna respuesta es correcta.

}

En redes neuronales se aplican una serie de operaciones para obtener funciones de activación derivables cuando se requiere aplicar el algoritmo de entrenamiento. Pero, ¿qué problema es el que nos lleva a utilizar precisamente este entrenamiento multicapa?: {

- ~ El problema del overshooting
- ~ El problema de los mínimos locales
- = El problema de la No-separabilidad lineal

}

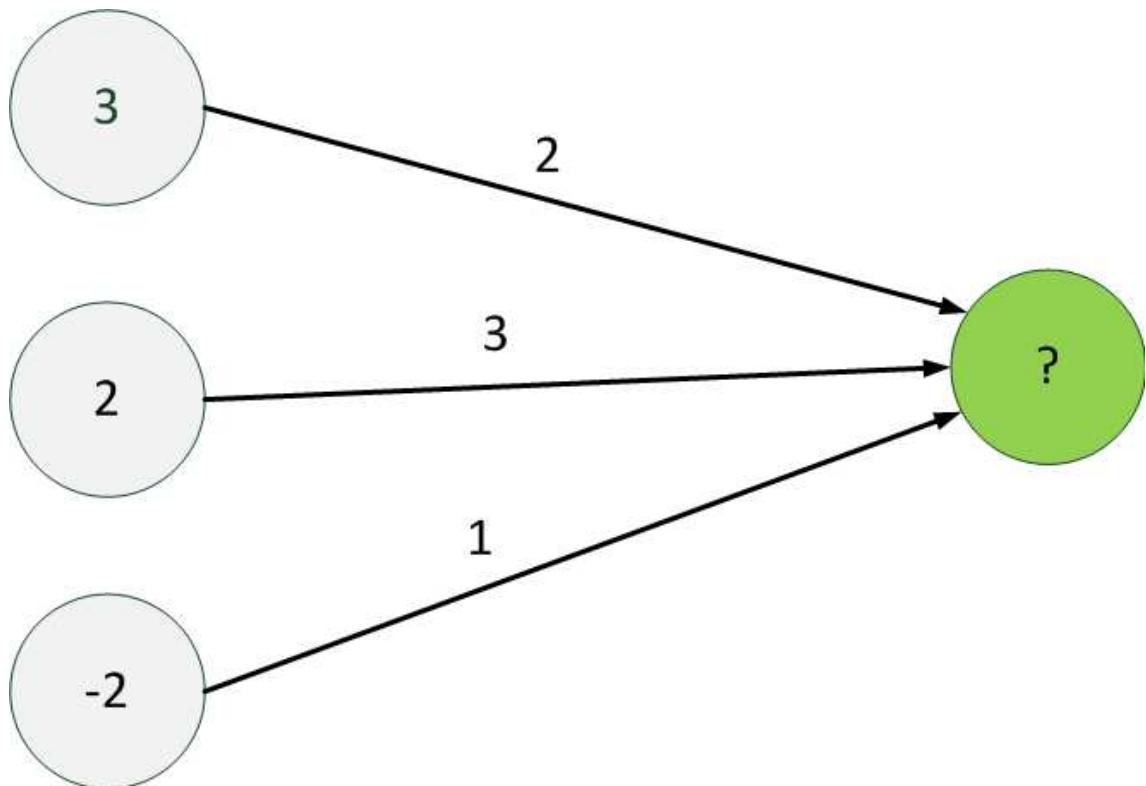
¿Qué afirmación sobre el aprendizaje por la Regla Delta es correcta? {

- ~ Utiliza el método de aprendizaje No Supervisado
  - ~ Utiliza el método de aprendizaje Razonamiento inductivo
- = Utiliza el método de aprendizaje Supervisado
- }

Si el resultado obtenido usando la Regla Delta es demasiado grande comparado con la salida deseada... {

- ~ Se aumenta el peso de las aristas.
  - ~ Se hace nada.
- = Se reduce el peso de las aristas.

}



{Qué **netinput** recibe esta neurona (verde)?{

~ 4

~ 12

= 10

}

Tenemos una red neuronal artificial compuesta por una única neurona con dos entradas.

Los pesos de estas entradas son  $w_1 = 3$ ,  $w_2 = -2$  y la función de activación o umbral de la neurona es  $\theta = 2$ .

De los siguientes elementos [X = (x1, x2)]: A = (3, -2); B = (2, 3); C = (-1, 4); D = (3, 2).

¿Cuáles activarían la salida de la neurona si la función de transferencia devuelve '1' solo para valores de entrada mayores que '0'?

{

~ A, B y C

~B, C y D

=A, B y D

}

En las redes neuronales, respecto a la regla Delta, señala el enunciado correcto:{

- ~ Se asume que el decremento de los pesos es proporcional a la disparidad entre la entrada observada y la salida observada.
- ~ Se asume que el incremento de los pesos es proporcional a la disparidad entre la salida observada y la entrada deseada.
- = Se asume que el incremento de los pesos es proporcional a la disparidad entre la salida observada y la salida deseada.

}

Qué afirmación es correcta sobre la regla delta (entrenamientos):{

- ~ El incremento de los pesos es inversamente proporcional a la disparidad entre la salida observada y la salida deseada.
- ~ La proporcionalidad de la disparidad entre la salida observada y la salida deseada no viene modulada por la constante de aprendizaje.
- = El incremento de los pesos es proporcional a la disparidad entre la salida observada y la salida deseada.

}

¿Cuál es el principal problema de la inicialización de la red en los descensos por gradiente?:{

- ~ Se obtienen máximos locales.
- ~ Se obtienen mínimos locales pero lo que se realizan ascensos por gradiente y no descensos.
- = Se obtienen mínimos locales.

}

Respecto a las fases del Backpropagation podemos decir que: {

- = Consta de varias fases, hacia delante, hacia atrás y la actualización de los pesos de todas las capas.
- ~ Consta de una única fase en la que se actualizan los pesos de todas las capas.
- ~ Ninguna de las anteriores es correcta.

}

La regla delta: {

- = Permite ajustar iterativamente el plano.
  - ~ Permite ajustar recursivamente el plano.
  - ~ Ninguna de las anteriores es correcta.
- }

En el entrenamiento de una red neuronal tenemos que  $\omega_i = \omega_i + \eta (d - y) x_i$ . Si la constante de aprendizaje  $\eta = 1$ ,  $x_i = 2,5$  y en la iteración anterior teníamos que  $\omega_i = 1$ , ¿cuáles son los valores correctos para esta iteración de  $\omega_i$ ? {

- = Un posible valor es  $\omega_i = -1.5$
- ~ No es posible saberlo sin los valores de 'd' e 'y'
- ~  $\omega_i$  no puede tomar valores reales, será un valor acotado entre 0 y 1}

Una red neuronal converge cuando{

- = El error de validación se mantiene bajo y los ejemplos de entrenamiento no provocan cambios significativos en los pesos de la red.
- ~ El error de validación se mantiene bajo y los ejemplos de entrenamiento provocan cambios significativos en los pesos de la red.
- ~ El error de validación se mantiene alto y los ejemplos de entrenamiento provocan cambios significativos en los pesos de la red.

}

Di cual es falsa de estas afirmaciones

{

~Usaremos más de una neurona en un problema cuando un único hiperplano no puede separar los datos

=A la capa adicional dentro de los Perceptrones multi-capa se le denomina capa visible

~ Una red neuronal converge cuando el error de validación se mantiene bajo y los ejemplos de entrenamiento no provocan cambios significativos en los pesos de la red.

}

¿Cómo se realiza la integración en una neurona computacional? {

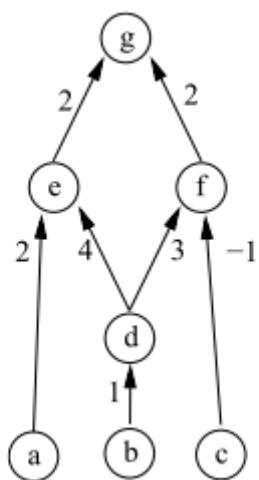
~ suma ponderada (net) por los pesos sinápticos seguida de varias funciones de activación  $f(\text{net})$ .

~ multiplicación ponderada (net) por los pesos sinápticos seguida de una función de activación  $f(\text{net})$

= suma ponderada (net) por los pesos sinápticos seguida de una función de activación  $f(\text{net})$

}

Dado una función de activación ( $y_i = 2 \cdot \text{in}_i = 2 \cdot \sum_j w_{ji} x_j$ ; donde  $\text{in}_i$  significa "input para neurona  $i$ ") y la siguiente red:



Cual de los siguientes valores es el resultado correcto de la salida de la neurona  $g$  ( $\rightarrow y_g$ ) para un vector de entrada  $(a, b, c) = (-1, 1, -1)$ ?

{

= 104

~ 98

~ 107

}

Las redes neuronales se basan en el tipo de aprendizaje: {

~ Por refuerzo

~ No Supervisado

= Supervisado

}

Indica la opción correcta acerca del algoritmo Backpropagation para redes neuronales

{

~ El algoritmo Backpropagation hace iteraciones hasta que el error supere un umbral.

~ El algoritmo Backpropagation hace iteraciones hasta que el error iguale un umbral.

= El algoritmo Backpropagation hace iteraciones hasta que el error baje de un umbral.

}

Cúal de las siguientes afirmaciones acerca de las reglas de aprendizaje neuronal NO es correcta?

{

~ las reglas *Delta* y *Backpropagation* ambas son reglas para aprendizaje supervisado.

~ *Backpropagation* se puede aplicar a redes con capas oculta pero la regla *Delta* no.

= Las fases de la regla *Delta* son: 1. Hacia delante, 2. Hacia atrás, 3. actualización de los pesos de todas las capas.

}

Una neurona computacional divide en 3 partes que son:{

~ Entradas, desintegración y Salida.

= Entradas, Integración y Salida.

~ Entradas, procesamiento y Salida.

}

La interpretación geométrica de la función de activación "ubral" es:{

= Un hiperplano .

~ No es posible hacer una interpretación geométrica.

~ Ninguna de las anteriores.

}

¿Como podemos detectar el sobreentrenamiento?

{

- ~ Cuando el error de entrenamiento sube despues de haber alcanzado niveles minimos.
  - = Cuando el error de validación sube despues de haber alcanzado niveles minimos.
  - ~ Ambas son correctas.
- }

{En qué consiste el sobreaprendizaje?: {

- ~ En que el sistema sea capaz de responder adecuadamente tras la convergencia .
- ~ En que el sistema sea incapaz de responder adecuadamente por no encontrar convergencia durante el aprendizaje.
- = Ninguna de las anteriores.

}

Los pesos de la red se pueden inicializar: {

- = Arbitraria y aleatoriamente.
  - ~ Arbitraria y paralelamente.
  - ~ Arbitraria y secuencialmente.
- }

Una red neuronal converge cuando... {

- ~ el error de validación se mantiene alto.
- ~ el error de validación alcanza valores muy pequeños.
- = el error de validación se mantiene bajo.

}

Según el modelo de una neurona computacional y una biológica, relaciona los siguientes términos: dendritas,soma,axon,suma ponderada,nºreales y resultado:{

- ~ Dendritas, resultado; Axón, nº reales; Soma, suma ponderada. Siendo las primeras de cada pareja las partes de una neurona computacional y la segunda las de una neurona biológica.
- ~ Dendritas, resultado; Axón, nº reales; Soma, suma ponderada. Siendo las primeras de cada pareja las partes de una neurona biológica y la segunda las de una neurona computacional.

= Dendritas, nº reales; Axón, suma ponderada; Soma, resultado. Siendo las primeras de cada pareja las partes de una neurona biológica y la segunda las de una neurona computacional.

}

La convergencia de backpropagation para una red neuronal: {

~ Busca el peso más bajo.

~ Provoca cambios significativos en los pesos de la red.

= Converge cuando el error de validación se mantiene bajo.

}

El proceso de aprendizaje de las neuronas se realiza cuando {

= hay varios cambios significativos en las sinapsis de las neuronas.

~ No se aprende. Ya empeza capacitado.

~ ninguna de las anteriores.

}

El proceso de funcionamiento de una neurona computacional es... {

= Recibe una cantidad de números, hace una suma ponderada por los pesos sinápticos, elabora una función de activación y obtiene un resultado a partir de esta.

~ Recibe un único dato, crea diversas funciones de activación a partir de este dato y obtiene un resultado a partir de la suma de esas funciones.

~ Recibe una cantidad de números, hace una suma ponderada por los pesos sinápticos, elabora diversas funciones de activación a partir de la suma y obtiene un resultado para cada una.

}

Acerca de la regla delta, no es correcta:{

~ Permite ajustar iterativamente el hiperplano.

= Se asume que el decremento de los pesos es proporcional a la disparidad entre la salida observada y la salida deseada.

~ Dicha proporcionalidad viene modulada por la constante de aprendizaje.

}

Centrándonos en Backpropagation: derivación matemática{

- = El algoritmo es un descenso por gradiente
  - ~ El algoritmo es un ascenso por gradiente.
  - ~ Hay que modificar los  $w_j$  en la dirección del gradiente.
- }

En cuanto a los preceptores multicapa... {

- ~ No presenta problemas con regiones de decisión más complejas exigen distintas estrategias de separación.
  - ~ La capa de regiones convexas se denomina capa oculta.
- = Se demuestra que un perceptrón con dos capas ocultas puede aproximar cualquier función

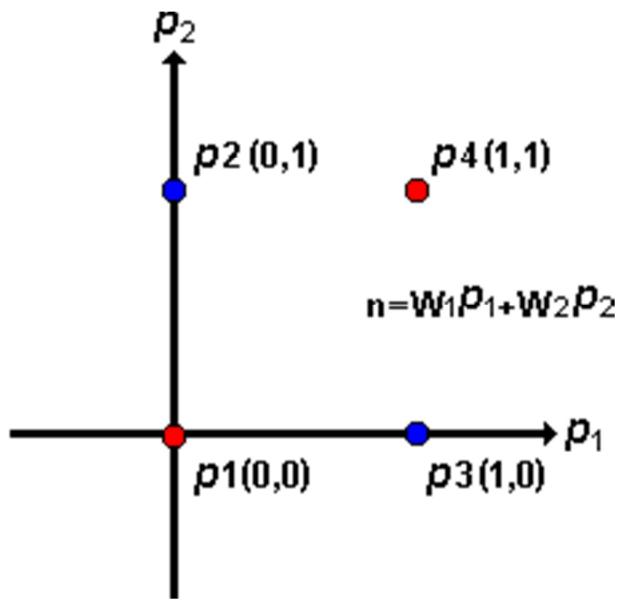
}

Pregunta: ¿Qué hace falta para que una neurona dispare salida? {

- ~ Siempre que estén todas las entradas posibles ocupadas.
- = Si las entradas superan un umbral.
- ~ Siempre que haya, al menos, una entrada.

}

En la siguiente figura se muestra el plano formado por el **problema del XOR**.



Donde la **clase 1** son los puntos rojos y la **clase 2** los puntos azules, y teniendo en cuenta que para los valores de entrada 00 y 11 se debe devolver la clase 0 y para los patrones 01 y 10 la clase 1.

El problema se resuelve: {

= Utilizando dos perceptrones.

~ Utilizando un perceptrón.

~ Sin necesidad de utilizar ningún perceptrón.

}

En el entrenamiento de los perceptrones de una neurona, 'sin capas ocultas', cuando utilizamos la Regla Delta, si después de transcurrir cierto número de iteraciones, todos los ejemplos están bien etiquetados, diremos que tenemos: {

~ No-separabilidad lineal.

= Conjuntos de ejemplo linealmente separables.

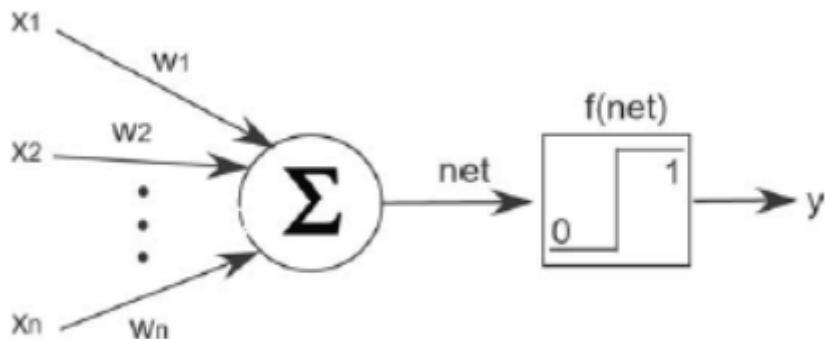
~ Una capa oculta.

}

Cual de las siguientes afirmaciones no es correcta {

~ En una neurona computacional las entradas son numeros reales

- ~ En una neurona biologica las entradas son las dendritas
- = La salida de una neurona biologica se transporta por el soma
- }



Acerca de las neuronas artificiales como la de la figura, indica que afirmación es correcta:

- {
- ~ Las entradas  $X_1, X_2, \dots, X_n$  pueden ser cualquier tipo de número.
- =  $W_1, W_2, \dots, W_n$  son los pesos sinápticos y determinan la influencia de cada entrada en la activación de la neurona, siendo excitatoria si  $W_i$  es positivo o inhibitoria si  $W_i$  es negativo.
- ~ La suma de todas las entradas ponderadas es el valor de salida de la neurona.
- }

La inicialización de los pesos de una red neuronal es:{

- ~ Arbitraria y segura y no supone ningún problema.
- = Arbitraria y aleatoria y el problema de la inicialización en los descensos por gradiente si tiene solución.
- ~ Arbitraria y aleatoria y el problema de la inicialización en los descensos por gradiente no tiene solución.
- }

¿Cuál puede ser una posible solución a la inicialización de los pesos de la red de una forma aleatoria o arbitraria?{

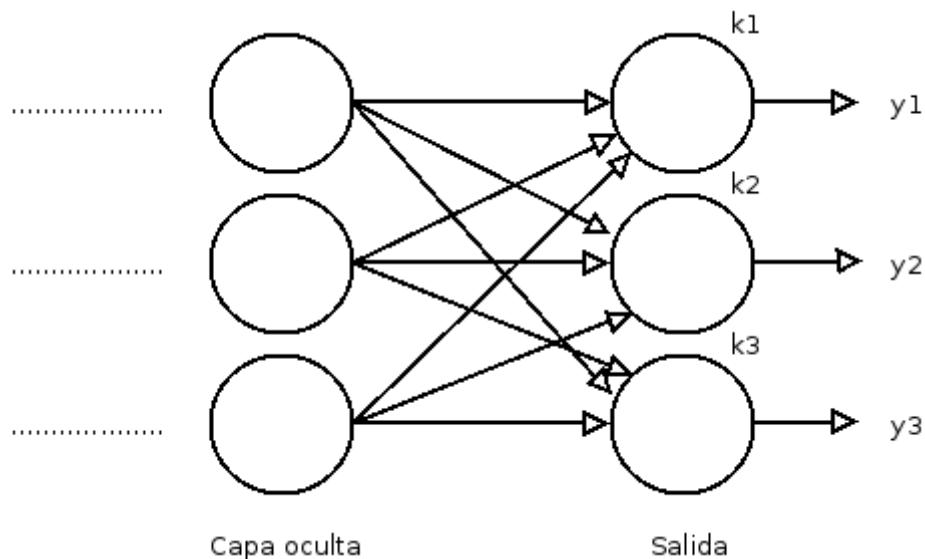
- ~ Ninguna, no hace falta hacer nada más

= Entrenar la red desde distintas inicializaciones

~ Encontrar el mínimo local

}

Dada la siguiente imagen y teniendo en cuenta la siguiente formula, elige la opción correcta:



$$\delta_k = (d_k - y_k) f'(net_k),$$

{

= El valor de  $\delta_{k1}$  es 1, el valor de  $\delta_{k2}$  es 1 y el valor de  $\delta_{k3}$  es 0, eso implica que la neurona k1 y k2 son responsables, cada uno, de tener un error.

~ El valor de  $\delta_{k1}$  es 1, el valor de  $\delta_{k2}$  es 1 y el valor de  $\delta_{k3}$  es 0, eso implica que la neurona k3 es responsable de un error.

~ Ninguna neurona da error.

}

{Cual de las siguientes afirmaciones es correcta sobre clasificadores débiles?: {

- = Son moderadamente precisos, simples y funcionan al menos mejor que una clasificación aleatoria.
- ~ Son altamente precisos, complejos y funcionan mejor que una clasificación aleatoria.
- ~ Ninguna de las anteriores es correcta.

}

Sobre los clasificadores supervisados, estan compuestos por:{

- = Patrones, Clases, Conjunto de entrenamiento, Función aprendida, Clasificador.
- ~ Patrones, Clases, Conjunto de entrenamiento, Función de evaluación, Clasificador.
- ~ Ninguna de las anteriores.

}

Indica cuál de las siguientes afirmaciones es INCORRECTA acerca de la notación de "AdaBoost (Adaptive Boosting)":

{

- ~ " $D_t$ " es dependiente respecto a la complejidad de los ejemplos.
- = " $i$ " indexa clasificadores (débiles), mientras que " $t$ " indexa ejemplos.
- ~ " $Z_t$ " es una constante de normalización.

}

En cuanto al Boosting y Bagging{

- ~ Bagging utiliza un voto ponderado
- ~ Boosting produce un clasificador más débil
- = Boosting NO combina los clasificadores con el mismo peso en el voto.

}

Usando Adaboost, cuando actualizamos la Distribución D nos encontramos con que:{

- = Inicialmente, cuando  $T=1$  todos los ejemplos son igualmente probables.
- ~ A lo largo de todo el proceso de Adaboost nos vamos a encontrar con que todos los ejemplos son igualmente probable.

~ Adaboost es un proceso que inicialmente determina una probabilidad por defecto de nula. Al principio ningun ejemplo es probable.

}

En Adaboost, el valor de  $\alpha_t$  surge de intentar optimizar el error asociado a  $h_t$ ,  $\epsilon_t$ , y es: {

$$\sim \alpha_t = 1/2 \ln (1 + \epsilon_t) / \epsilon_t$$

$$= \alpha_t = 1/2 \ln (1 - \epsilon_t) / \epsilon_t$$

$$\sim \alpha_t = 1/2 \log_2 (1 - \epsilon_t) / \epsilon_t$$

}

Adaboost es un algoritmo utilizado para construir.....{

= clasificadores sólidos utilizando la combinación lineal de clasificadores simples

~ clasificadores débiles.

~ mediante Bagging clasificadores débiles

}

Selecciona la respuesta incorrecta acerca de Boosting y Bagging: {

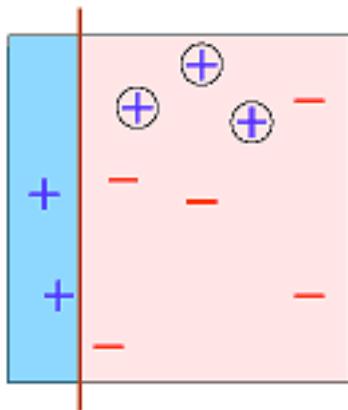
~ Se usa un voto ponderado.

= Se combinan los clasificadores con el mismo peso en el voto.

~ Ambas son incorrectas.

}

Dados los siguientes datos de una iteración usando el algoritmo Adaboost:



¿A que muestras se les asignará una mayor ponderación? = {

- ~ A situadas a la izquierda del clasificador (zona azul).
  - = A las redondeadas.
  - ~ A las que no están redondeadas, puesto que están bien clasificadas.
- }

En el algoritmo Adaboost se persigue: {

- = Mantener un peso en cada uno de los ejemplos de entrenamiento.
  - ~ Mejorar los clasificadores estables.
  - ~ Que los ejemplos más cercanos a la frontera de decisión son más fáciles de clasificar, y recibirán pesos más bajos.
- }

En la siguiente iteración del ejemplo visto en clase,

dado que los valores de cada elemento es de 0.1, ¿Cuál será el valor de la ganancia?

- {
- ~ 0'3
  - ~ 0,64
  - = 0'42
- }

{Cúal de las siguientes características sobre el clasificador boosting es correcta?: {

- ~ Cada modelo individual se induce de manera separada.
- = Cada nuevo modelo está influenciado por el comportamiento de los anteriores.
- ~ Los nuevos modelos se convierten en inexpertos para ejemplos mal clasificados por los anteriores modelos.

}

Selecciona la afirmación correcta: {

- ~ Bagging es un algoritmo mejorado de Boosting.
- = Ejemplos de clasificadores inestables son redes neuronales y árboles de decisión.
- ~ Con Boosting se usan votos no ponderados.

}

D<sub>3</sub>

-	-	+	+	+	-	-	-	+	-
0.05	0.05	0.18	0.18	0.15	0.08	0.05	0.05	0.05	0.05

FALLO FALLO FALLO

Dada este ejemplo de BOOSTING cual es su  $\epsilon$ ,

{

~ 0,71

~ 0,12

= 0,18

}

Explicación: El épsilon es la suma de los errores en esa frontera.  $0.05 + 0.05 + 0.08 = 0.18$ .

En una cierta iteración del algoritmo AdaBoost, obtenemos  $\epsilon_t = 0$ . ¿Qué podemos determinar? {

- ~ El clasificador se ha equivocado en la iteración anterior.

- ~ El clasificador ha establecido una frontera errónea.
  - = El clasificador ha establecido una frontera perfecta.
- }

Indica, sobre el algoritmo de AdaBoost, cuál es la opción correcta: {

- ~ El valor de confianza depende del error que se cometió en el vector de pesos.
  - = El valor de confianza depende del error que se cometió en la clasificación débil.
  - ~ El valor de confianza depende del error que se cometió en la normalización.
- }

En la fórmula correspondiente al algoritmo Adaboost...

$$D_{t+1}(i) = \frac{D_t(i) \cdot e^{-\alpha_t \cdot y_i \cdot h_t(x_i)}}{Z_t}$$

¿con qué se corresponde el valor  $y_i$ ?

- {
- ~ Con el vector de pesos.
  - = Con la clase a la que pertenece el ejemplo.
  - ~ Con el subíndice de entrenamiento.
- }

Sobre el algoritmo de implementación de Bagging (Bootstrap aggregating) indica que afirmación es correcta:

{

~ Los elementos del conjunto de entrenamiento clasificados por  $h(t)$  no se pueden usar en  $h(t+1)$ .

~ En la hipótesis final se selecciona el clasificador  $h(t)$  que mejor haya evaluado el conjunto de entrenamiento.

= Ninguna de las anteriores es correcta.

}

En relación con Boosting y Adaboost, teniendo en cuenta la fórmula de actualización del algoritmo Adaboost

$$D_{t+1}(i) = \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

Selecciona la respuesta correcta:{

~ La variable  $(i)$  indexa clasificadores (débiles).  $(t)$  indexa ejemplos y  $\alpha_t$  es una constante de normalización.

~ La variable  $(i)$  indexa ejemplos.  $(t)$  indexa clasificadores (débiles) y  $\alpha_t$  es el error asociado a  $(i)$ .

= La variable  $(i)$  indexa ejemplos.  $(t)$  indexa clasificadores (débiles) y  $\alpha_t$  es una constante de normalización.

}

Cual de las siguientes formulas corresponde al error asociado a  $h_t$ : {

~  $\varepsilon_t = \Pr_{Dt}[h_t(x_i) = y_i]$

=  $\varepsilon_t = \Pr_{Dt}[h_t(x_i) \neq y_i]$

~  $\varepsilon_t = \Pr_{Dt}[h_t(x_i) = x_i]$

}

{Cuál de las siguientes características sobre Bagging es correcta?{

- = Ayuda a mejorar clasificadores inestables
- ~ Ayuda a mejorar clasificadores estables
- ~ Ninguna de las anteriores

}

Comparando Boosting y Bagging{

- ~ Boosting pondera y da más peso a los ejemplos que más cuestan clasificar, así en las siguientes iteraciones los clasificadores se centren en clasificar aquellos con más peso.
- ~ Bagging: entrena un clasificador débil con el subconjunto cogido T veces, por lo que obtendremos T clasificadores entrenados.
- = Ambas son correctas

}

{Qué afirmación acerca de Adaboost es falsa?

$$D_{t+1}(i) = \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

{

- ~ La formula de la imagen, sirve para actualizar la distribución D.
- ~  $\alpha_t$  depende del error  $\epsilon_t$  asociado a la  $h_t$
- =  $Z_t$  no es constante

}

Cúal de las siguientes afirmaciones es correcta:{

- ~ Tanto boosting como bagging son clasificadores débiles
- = Tanto boosting como bagging mejoran clasificadores inestables como por ejemplo las redes neuronales
- ~ Ninguna de las anteriores es correcta

}

En el Boosting, ¿hasta qué punto se entrena? {

- ~ Hasta que el valor de confianza sea cero.
  - = Hasta que se consigue clasificar bien el máximo de ejemplos posibles.
  - ~ Ninguna de las anteriores es correcta.
- }

Adaboost es un algoritmo utilizado para: {

- = construir clasificadores sólidos utilizando la combinación lineal de clasificadores simples.
  - ~ construir clasificadores simples utilizando la combinación lineal de clasificadores sólidos.
  - ~ construir clasificadores simples utilizando la combinación lineal de clasificadores inestables.
- }

Respecto a la implementación de boosting, selecciona la respuesta correcta:{

= Adaboost es la técnica que lo implementa, y el proceso que lleva a cabo es:

1. Entrenar un clasificador débil usando  $D_t$  y obtener  $h_t$
2. Escoger un valor de confianza  $\alpha_t$
3. Actualizar la distribución  $D$

~ Bagging es la técnica que lo implementa, y el proceso que lleva a cabo es:

1. Entrenar un clasificador débil usando  $D_t$  y obtener  $h_t$
2. Escoger un valor de confianza  $\alpha_t$
3. Actualizar la distribución  $D$

~ Adaboost es la técnica que lo implementa, y el proceso que lleva a cabo es:

1. Entrenar un clasificador débil usando  $D_t$  y obtener  $h_t$
2. Obtener la suma de los pesos mal clasificados
3. Actualizar la distribución  $D$

}

¿Qué afirmación sobre el método *Bagging* es correcta?{

- ~ Utiliza votos ponderados para la combinación de los clasificadores débiles.
- ~ Se hace una muestra aleatoria de los datos de entrenamiento *sin sustitución*.
- = Los modelos o clasificadores tienen los mismos pesos en la formación de la hipótesis final.
- }

Para construir y usar un Dt{

- = primero entrenar un clasificador débil usando Dt y obtener ht, segundo, escoger un valor de confianza ct y por último actualizar la distribución D.
- ~ primero eliminamos todos los clasificadores débiles quedandonos con el fuerte, segundo actualizamos la distribución D.
- ~ primero cogemos todos los clasificadores débiles y los comparamos hasta encontrar el mas fuerte, segundo escogemos el que mas cerca se haya quedado del fuerte, tercero actualizamos la distribucion D.
- }

¿Cual de las siguientes afirmaciones no es correcta ?{

- ~ Cuando añadimos muchos clasificadores podemos estar memorizando los datos
- = A mas clasificadores añadamos mejor aprenderemos a clasificar los datos
- ~ Añadir clasificadores no siempre mejora la clasificación de los datos
- }

$$H(x) = \text{sign}(f(x)) = \text{sign} \left( \sum_{t=1}^T \alpha_t h_t(\mathbf{x}) \right)$$

La formula anterior pertenece a: {

- ~ Bagging
- ~ Adaboost
- = Ninguno de los anteriores. }

Respecto al muestreo ponderado...:{

- ~ Intuitivamente, los ejemplos más cercanos a la frontera de decisión son más fáciles de clasificar, y recibirán pesos más altos.
  - ~ Intuitivamente, los ejemplos más cercanos a la frontera de decisión son más fáciles de clasificar, y recibirán pesos más bajos.
  - = Ninguna de las anteriores es correcta.
- }

Tenemos 100 ejemplos a clasificar mediante el método de adaboost. El primer clasificador falla 5 de ellos al clasificarlos ¿Cuál será el peso de los acertados, una vez normalizados, para la siguiente iteración?

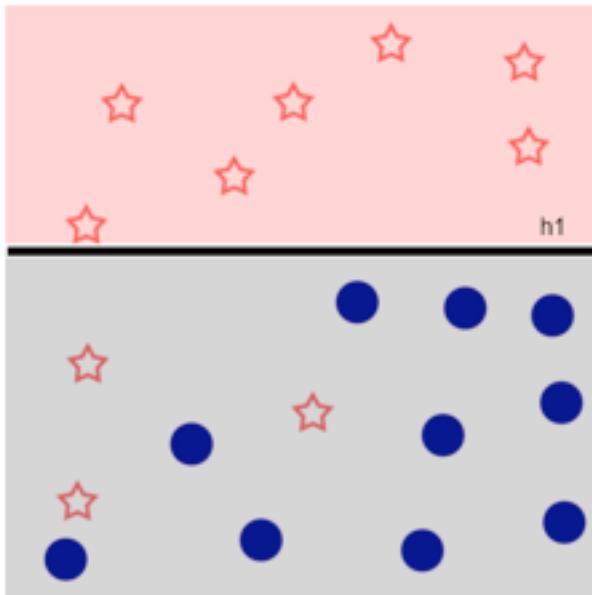
- {
  - = 0,0053
  - ~ 0,0105
  - ~ 0,0147
- }

En la comparación entre Boosting y Bagging. En cuanto al muestreo ponderado de Boosting, {

- ~ Se hace un muestreo aleatorio de los datos de entrenamiento.
- = Los ejemplos más cercanos a la frontera de decisión son los más difíciles de clasificar.
- ~ Se combinan los clasificadores con el mismo peso en el voto.

}

Dados los siguientes datos de la primera iteración usando el algoritmo Adaboost:



Indica el error  $\varepsilon_1$  asociado al clasificador  $h_1$  y el peso correspondiente  $\alpha_1$  usando  $D_t$  para  $T=1$

(Inicialmente, cuando  $T=1$  todos los ejemplos son igualmente probables así que  $D_1(i) = 1/N$  y  $\alpha_i = \frac{1}{2} \ln(1 - \varepsilon_i/\varepsilon_0)$ )

{

~  $\varepsilon_1 = 0,3, \alpha_1 = 0,42$

~  $\varepsilon_1 = 0,15, \alpha_1 = 0,79$

=  $\varepsilon_1 = 0,15, \alpha_1 = 0,87$

}

¿Por qué decimos que AdaBoost es un algoritmo adaptivo? {

~ Porque se puede combinar con muchos otros algoritmos de aprendizaje automático.

= Porque los subsecuentes clasificadores son construidos y mejorados en favor de los clasificadores anteriores mal clasificados.

~ Ninguna de las anteriores.

}

Con respecto a las características de Boosting y Bagging, cual es incorrecta:{

~ Boosting determina pesos para cada dato de entrenamiento;

- ~ Bagging hace un muestreo aleatorio de los datos de entrenamiento;
  - = Boosting concede mayor peso a las muestras que están clasificados correctamente y el más bajo a los mal clasificados.
- }

Para construir y usar Dt hay que seguir los siguientes pasos {

- = Primero entrenar un clasificador débil usando Dt y obtener ht, segundo escoger un valor de confianza at y último actualizar la distribución D.
- = Primero entrenar un clasificador débil usando Dt y obtener at, segundo escoger un valor de confianza ht y último actualizar la distribución D.
- = Primero escoger un valor de confianza at, segundo actualizar la distribución D y último entrenar un clasificador débil usando Dt y obtener ht.

}

Qué opción es INCORRECTA{

- ~ En el muestreo ponderado se ponderan las muestras para concretar el aprendizaje en los ejemplos difíciles.
- ~ En el voto ponderado, en lugar de combinar los clasificadores con el mismo peso en el voto, se utilizan votos ponderados.
- = En el voto ponderado, los ejemplos más cercanos a la frontera de decisión son más difíciles de clasificar, por lo que el peso de estos ejemplos será más alto.

}

¿Cuál de lo siguientes pasos son necesarios para realizar ADABoost: {

- ~ Calcula error del modelo en el set de entrenamiento
  - ~ Ponderar todos los sets de entrenamiento de igual forma
  - = Toas las anteriores son correctas
- }

En el algoritmo Adaboost, respecto al elemento H:{

- ~ Nunca debería ser más pequeño de 0,5.
- ~ Es la combinación de los clasificadores débiles, y por tanto también lo es.
- = Es un clasificador fuerte obtenido con la combinación de los clasificadores débiles entrenados.

}

Selecciona la respuesta correcta: {

- ~ En bagging los votos son ponderados unos mejores que otros.
- ~ En boosting los votos se ponderan de forma equitativa.
- = En boosting los votos son ponderados unos mejores que otros.

}

En el Boosting los ejemplos que reciben pesos más altos son:{

- ~ Los del centro, ya que tenemos más certeza de sus resultados.
- = Los ejemplos más cercanos a la frontera de decisión.
- ~ Ninguna de las anteriores.

}

los métodos de Boosting e Adaboost, son basados en

{

- =aprendizaje múltiple
- ~aprendizaje solo
- ~ambas son incorrectas

}

Si comparamos Boosting y Bagging: {

- ~ Bagging combina clasificadores débiles mientras que Boosting clasificadores fuertes.
- = Boosting realiza un muestreo ponderado mientras que Bagging no.
- ~ Ambas son correctas.

}

Los votos ponderados (clasificadores) en Boosting: {

- ~ Esta es la regla de combinación para el conjunto de clasificadores fuertes.

- = Esta es la regla de combinación para el conjunto de clasificadores débiles.
  - ~ Ninguna de las anteriores.
- }

{Cuál está definido como un clasificador débil? {

- ~ Árboles de decisión
- ~ Redes neuronales
- = Ambas respuestas son correctas.

}

Los pasos para construir y usar Dt son, en orden correcto;

- {
  - =1- Entrenar un clasificador débil, usando Dt y obtener Ht. 2- Escoger un valor de confianza αt. 3- Actualizar distribución D.
  - ~1- Escoger un valor de confianza αt. 2- Entrenar un clasificador débil, usando Dt y obtener Ht. 3- Actualizar distribución D.
  - ~1- Actualizar distribución D. 2- Entrenar un clasificador débil, usando Dt y obtener Ht. 3- Escoger un valor de confianza αt.
- }

Se puede probar que es posible encontrar un clasificador más preciso combinando muchos clasificadores "débiles". Para realizar dichas combinaciones, se pueden hacer mediante los métodos de: {

- ~ Boosting y Boinging.
  - ~ Boinging y Forwarding.
  - = Boosting y Bagging.
- }

Para que un clasificador sea considerado "debil" ha de cumplir:

- {
  - = Ser simple.
  - ~ Mejorar la clasificación de un perceptrón.
  - ~ Ambas son correctas.
- }

Pregunta: ¿Qué puede ocurrir si combinamos un conjunto de clasificadores? :{

- ~ Si son compatibles los clasificadores se obtendrá un clasificador más fuerte que estos pero si no lo son, el resultado será un clasificador más débil que estos por separado.
  - = Que obtendremos un clasificador más fuerte que estos por separado.
  - ~ Se obtendrá como resultado un clasificador más débil que estos por separado.
- }

¿Acerca de los votos ponderados (clasificadores), cual es la incorrecta? :{

- = Se usan los clasificadores con el mismo peso en el voto.
  - ~ Esta es la regla de combinación para el conjunto de clasificadores débiles.
  - ~ En conjunción con la estrategia de muestreo anterior, esto produce un clasificador más fuerte.
- }

Para Boosting y Adaboost, podemos afirmar que:{

- ~ Boosting es la técnica para entrenar varios clasificadores débiles para encontrar un clasificador mejor.
  - ~ Boosting podríamos decir que es el método que más se usa hoy en día porque se puede aplicar normalmente a cualquier problema y funciona extremadamente bien. Un ejemplo, el algoritmo de detección de caras de las cámaras.
  - = Las 2 opciones anteriores son correctas.
- }

Cuál es la diferencia entre Boosting y Bagging?{

- ~ el Boosting no asigna pesos a cada registro de entrenamiento y Bagging elige aleatoriamente los registros para formar los subconjuntos.
  - = el Boosting asigna pesos a cada registro de entrenamiento y Bagging elige aleatoriamente los registros para formar los subconjuntos.
  - ~el Boosting no asigna pesos a cada registro de entrenamiento y Bagging asigna ponderaciones a cada registro para formar los subconjuntos.
- }

Usando Dt; cuando T=1{

- ~ Eligiremos los que hacen fallar al clasificador
  - = Todos los ejemplos son igualmente probables
  - ~ Seleccionaremos los ejemplos más difíciles
- }

Indica cuál de las siguientes respuestas es correcta acerca del muestreo ponderado: {

- = Los ejemplos más cercanos a la frontera de decisión son más difíciles de clasificar, y recibirán pesos más altos.
  - ~ Los ejemplos más cercanos a la frontera de decisión son más difíciles de clasificar, y recibirán pesos más bajos.
  - ~ Los ejemplos más cercanos a la frontera de decisión son más fáciles de clasificar, y recibirán pesos más altos.
- }

Siguiendo la notación empleada en las transparencias de teoría, indica cual de estas respuestas es correcta:  
{

- ~ Bagging ayuda a mejorar clasificadores estables.
  - = En Adaboost entrenamos un clasificador débil usando  $D_t$  para obtener  $h_t$ .
  - ~ En Adaboost entrenamos un clasificador débil usando  $h_t$  para obtener  $D_t$ .
- }

¿Qué estrategia de muestreo produce un clasificador más fuerte?:{

- ~ Votos ponderados (Clasificadores).
  - ~ Muestreo ponderado (Ejemplos).
  - = La conjunción de las dos anteriores.
- }

Nos encontramos construyendo y usando  $D_t$ . A la hora de actualizar la distribución  $D$ :{

- ~ Siempre, cuando  $T=4$  todos los ejemplos son igualmente probables.
- ~ Inicialmente, cuando  $T=1$  los ejemplos no son igualmente probables.
- = Pasada la iteración  $T=1$ , en las siguientes es más probable seleccionar los ejemplos más difíciles (los que hacen fallar al clasificador).

}

Acerca de Bagging:{

= Ayuda a mejorar clasificadores inestables.

~ Grandes cambios en el conjunto de entrenamiento no producen grandes cambios en el porcentaje de aciertos.

~ Ambas son correctas.

}

Sobre Boosting ,Muestreo ponderado: {

= En lugar de hacer un muestreo aleatorio de los datos de entrenamiento, se ponderan las muestras para concentrar el aprendizaje en los ejemplos más difíciles.

~ Intuitivamente, los ejemplos más lejanos a la frontera de decisión son más difíciles de clasificar, y recibirán pesos más altos.

~ Ambas son correctas.

}

Bangging y Boosting:{

~ Son combinadores de clasificación débiles.

~ Son algoritmos de clasificación de combinadores débiles.

= Ninguna de las anteriores.

}

En cuanto a Boosting vs Bagging, selecciona la opción correcta{

~ Boosting puede dañar performance en datasets ruidosos.

~ En la práctica, Bagging casi siempre mejora el modelo.

= Las dos son correctas.

}

Cuando construimos y usamos  $D_t$  : {

~  $\alpha_t$  surge de intentar optimizar  $h_t$

=  $\epsilon_t$  es el error asociado a  $h_t$

~ Ambas son correctas

}

Construyendo y usando  $D_t$  (AdaBoost) al actualizar la distribución  $D$ , selecciona la opción correcta:{

~ Cuando  $T > 1$  todos los ejemplos son igualmente probables.

= En las siguientes iteraciones, es más probable seleccionar los ejemplos más difíciles.

~ Cuando  $T < 1$  todos los ejemplos son igualmente probables.

}

Sobre AdaBoost, ¿Cuál de las afirmaciones NO es correcta?:{

=  $Z_t$  varía en cada iteración de  $t = 1, \dots, T$ .

~  $D_t$  depende de la complejidad de los ejemplos.

~  $i$  indexa ejemplos, mientras que  $t$  indexa clasificadores.

}

En lo referente a Bagging, ayuda a mejorar clasificadores inestables como ...{

= Redes neuronales y árboles de decisión

~ Vectores y listas

~ Ninguna de las anteriores.

}

En Boosting, en el muestreo ponderado, los ejemplos más cercanos a la frontera de decisión: {

~ recibirán pesos más bajos.

= recibirán pesos más altos.

~ sus pesos no cambiaran.

}

Selecciona la respuesta correcta sobre las caracteristicas de Bagging:

{

=Ayuda a mejorar clasificadores inestables, como redes neuronales o árboles de decisión.Pequeños cambios en el conjunto de entrenamiento llevan a diferentes clasificadores y grandes cambios en el porcentaje de aciertos.

~ Ayuda a mejorar clasificadores estables, como redes neuronales o árboles de decisión.Grandes cambios en el conjunto de entrenamiento llevan a diferentes clasificadores y pequeños cambios en el porcentaje de aciertos.

~ Ninguna de las anteriores es correcta.

}

Los clasificadores débiles son moderadamente precisos. Si combinásemos **muchos** ¿Qué **podríamos** encontrar? {

= Un clasificador más preciso.

~ Un clasificador igual de preciso que los clasificadores combinados.

~Un clasificador más preciso y que además funciona peor que una clasificación aleatoria.

}

**El último "paso" del trabajo que realiza un detector de Canny es:**

{

= El algoritmo realiza la unión de los pixeles, incorporando “candidatos débiles” que están 8-conectados a los pixeles “probables”.

~ La imagen se suaviza usando un filtro Gaussiano con una desviación estándar para reducir el ruido. El gradiente local, y la dirección del borde, son computados en cada punto. El detector Sobel, puede ser usado para computar G<sub>x</sub> y G<sub>y</sub>.

~ Se define un punto de borde, como un punto cuyo peso es localmente máximo en la dirección del gradiente.

}

Explicación: El último punto de la forma en la que trabaja un detector de Canny es donde el algoritmo realiza la unión de los pixeles incorporando los pixeles que se encuentran en crestas con valores entre T<sub>1</sub> y T<sub>2</sub> que están 8-conectados a los pixeles de crestas con valores mayores que T<sub>2</sub>.

**¿Como podríamos saber en una imagen el numero de pixeles que comparten un mismo valor?**

{

= Con la función histograma

~ Aplicando convolución

~ Con el detector de Canny

}

Explicación: La función histograma definida en K niveles (valores de intensidad), indica el numero de pixeles en la imagen que contienen ese determinado valor.

**Con respecto al histograma de una imagen, se puede afirmar que el proceso de binarización:**

{

= Consiste en poner a 1 los pixeles que estén por encima de un umbral o entre dos umbrales y el resto a 0

~ No funciona con imágenes a color

~ Es necesario definir como mínimo 2 umbrales para poner a 1 los pixeles que se encuentren en medio de estos y el resto a 0

}

Explicación: La Binarización consiste en indicar uno o dos umbrales y poner a 1 los pixeles que estén por debajo o por encima del umbral o entre dos umbrales, y a cero el resto. Y se puede realizar en imágenes de color aplicando umbrales a cada banda de color.

**¿Qué ocurrirá si aumentamos el valor de todas las posiciones de una matriz de imagen en escala de grises?**

{

- = Aumentará el brillo de la imagen.
- ~ La imagen se volverá más oscura.
- ~ La imagen se difuminará.

}

Explicación: Debido a que los colores van desde 0 a 255, siendo el 0 el negro y el 255 el blanco, si aumentamos los valores se acercarán mas al blanco con lo que aumentará el brillo de la imagen.

**En los filtros sobre imágenes, ¿por qué hay que tener en cuenta que el tamaño de la máscara usada o de la vecindad escogida (Efecto del tamaño de la máscara) :**

{

- = Puede tener un efecto no deseado en el resultado, aunque hay veces que es lo que se pretende.
- ~ No afecta en absoluto a la calidad de la imagen.
- ~ Tiene una buena detección, maximiza el número de falsos positivos y minimiza los falsos negativos

}

Explicación: Hay que tener en cuenta que el tamaño de la máscara usada o de la vecindad escogida (3x3, 5x5, 10x10) puede tener un efecto no deseado en el resultado, aunque hay veces que es lo que se pretende. Viene en la página 10 del tema 11.

**¿Cuál de las siguientes afirmaciones es INCORRECTA acerca de los histogramas de la imagen?:**

{

- ~ Indica el número de píxeles en la imagen que tienen el valor de intensidad comprendido en una función definida en K niveles.
- = La ecualización del histograma consiste en comprimirlo para mejorar el contraste.
- ~ La binarización puede aplicarse, además de a imágenes en blanco y negro, a imágenes en color.

}

Explicación: Como bien viene documentado en las diapositivas 11,12 y 13 del tema 11 de la asignatura, la opción incorrecta es la segunda, dado que la ecualización del histograma se aplica a imágenes con poco contraste (para mejorarlo), pero el proceso consiste en expandir dicho histograma, no contraerlo.

**Cuando reducimos la resolución de una imagen, al perder píxeles:**

{

~ Promedia entre los píxeles adyacentes.

= Promedia entre los píxeles eliminados.

~ Ninguna de las anteriores.

}

Explicación: Como pone en la página 4 del tema 11 de teoría es una operación básica para reducir el tamaño de imagen el promediar entre los eliminados o elegir uno.

### Respecto a los criterios del Detector de Canny ¿Que implica una "Buena detección"?

{

~ Las aristas deben marcarse en el lugar real.

= Se debe minimizar el número de falsos positivos y falsos negativos.

~ Como resultado se deben generar aristas de un pixel de anchura.

}

Explicación: Como podemos observar en la diapositiva nº21 del tema 11 y en diversas preguntas de los compañeros. Los criterios del detector de canny son:

- Buena detección

- Buena localización

- Respuesta única

Pero, ¿en este contexto que debe cumplir cada una de ellas? una buena detección sería la segunda respuesta (marcada como correcta), buena localización sería la primera, y una respuesta única la tercera respuesta.

### Cuál es el filtro usado para suavizar la imagen en el Detector de Canny?

{

= Filtro Gaussiano

~ Filtro de Mediana

~ Ninguna de las anteriores

}

Explicación: En el Detector de Canny la imagen es suavizada usando un filtro Gaussiano con una desviación estándar, esto para reducir el ruido.

### En cuanto a los modelos de color, el tipo de imagen que tendremos en función de la matriz puede ser

{

- ~ Gris, valores -1, 0 y 1
  - ~ RGB (Red, Green, Black), los tres valores entre 0 y 255
  - = Binaria, valores 0 y 1.
- }

Explicación: La solución es la c), el modelo a) no existe y el modelo RGB se basa en los colores Rojo, verde y azul, no en el negro.

### Sobre el detector de Canny:

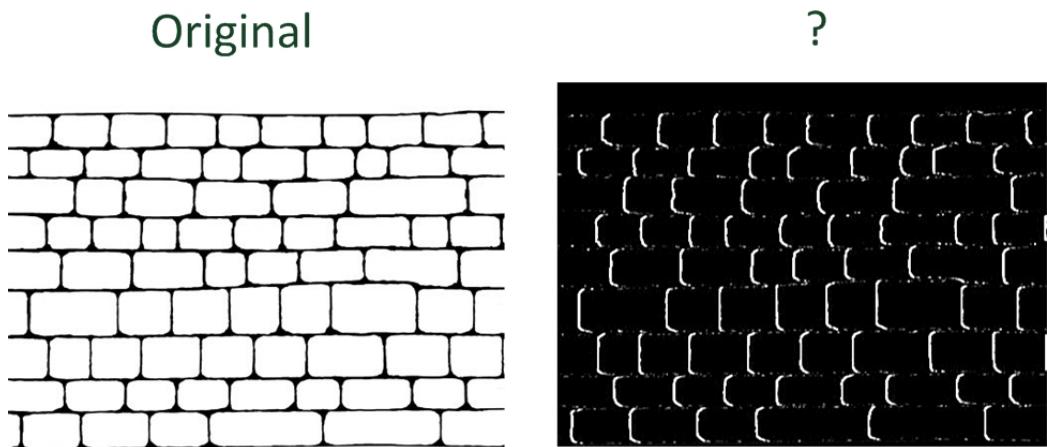
- {
- ~ es el más efectivo a la hora de detectar esquinas.
  - ~ es el más efectivo a la hora de detectar bordes, debido a su eficiencia.
  - = es el más efectivo a la hora de detectar bordes, debido a su eficacia.
- }

Explicación: El detector Canny es considerado, el más efectivo a la hora de detectar bordes, debido a su eficacia. (Transparencia nº23)

**La ecualización de un histograma es una operación básica en el tratamiento de imágenes que resulta útil en la percepción automática. Para realizar dicha operación es necesario seguir un algoritmo que consiste en una serie de pasos. Señala la respuesta correcta:**

- {
- = Paso 1: Se calcula el histograma de la imagen  $h(k)$ . Paso 2: Se calcula el histograma acumulado  $H_a(k) = \sum$  (desde  $j=0$  hasta  $k$ )  $h(j)$ . Paso 3: se produce un mapeo con la siguiente fórmula  $I'(x,y) = (H_a(I(x,y)) - minv) * (L - 1)/(totalpix-minv)$
  - ~ Paso 1: Se calcula el histograma de la imagen  $h(k)$ . Paso 2: Se calcula el histograma acumulado  $H_a(k) = \sum$  (desde  $j=0$  hasta  $k$ )  $h(j)$ . Paso 3: se produce un mapeo con la siguiente fórmula  $I'(x,y) = (H_a(I(x,y)) - minv) * (minv - 1)/(totalpix-L)$
  - ~ Paso 1: Se calcula el histograma de la imagen  $h(k)$ . Paso 2: Se calcula el histograma acumulado  $H_a(k) = \sum$  (desde  $j=0$  hasta  $k$ )  $h(j)$ . Paso 3: se produce un mapeo con la siguiente fórmula  $I'(x,y) = (H_a(I(x,y)) - minv) * (L - 1)/(totalpix-maxL)$
- }

Explicación: Para calcular la ecualización del histograma son necesarios tres pasos, primeramente calculamos el histograma de la imagen, posteriormente calculamos el histograma acumulado y finalmente utilizamos la fórmula del mapeo. Donde L es el número de niveles de intensidad y minv es el valor de ha correspondiente al menor valor de gris. Como se puede ver en las respuestas y concretamente en la fórmula de mapeo vemos que maxL es un parámetro inventado.



**¿Qué operador se utilizó para obtener la segunda imagen?**

{

~ Operador Sobel

$$\mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

~ Filtro de media

$$\frac{1}{9} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

= Operador Sobel

$$\mathbf{G}_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix}$$

}

Explicación:

Para detectar aristas en imágenes se pueden utilizar los operadores Sobel. En este ejemplo se utilizó el operador  $G_x$  para detectar las aristas verticales. Para detectar las aristas horizontales se puede utilizar el operador  $G_y$  que produciría esta imagen inferior:

**¿Cuál es la imagen resultante de aplicar un filtro de mediana de tamaño 3x3 a la siguiente imagen?**

20 23 30 31 {

22 24 28 30

23 26 30 33

= a

~ b

~ c

}

Explicación: Para aplicar el filtro de mediana se sustituye el valor del píxel estudiado por la mediana de los valores que engloba una ventana de selección dada. La mediana será el píxel que se encuentre en la posición central al ordenarlos según valor.

En este caso los píxeles que se sustituyen son 21, 29, 24 y 32.

El cálculo de las medianas es el siguiente:

21 -> 20 21 22 23 23 24 29 30 32 = 23

29 -> 21 23 24 29 30 30 31 33 33 = 30

24 -> 21 22 23 24 29 29 31 32 34 = 29

32 -> 21 24 29 30 31 32 33 34 37 = 31

El resultado es 23, 30, 29 y 31, que corresponde a la respuesta a.

**Dado un histograma con los siguientes datos:**

**K-Nivel h(k)**

1	458
2	321
3	276
4	230
5	124
6	57

**¿Cuál será el resultado de ecualizar el histograma?**

{

= a

Nivel      N° Pixeles

1	458
2	0
3	321
4	276
5	230
6	181

~ b

Nivel      N° Pixeles

1	221
2	237
3	321
4	276
5	230
6	181

~ c

Nivel      N° Pixeles

1	458
2	127
3	192
4	278
5	210
6	201

}

Explicación: De la tabla del enunciado podemos deducir los siguientes datos:

minv	458
total	1466
L	6

Con estos datos y la tabla anterior, aplicamos la fórmula  $I'(x,y) = (ha(I(x,y)) - minv) / (total - minv) * (L-1)$ , después de haber calculado la distribución acumulada de los píxeles.

Obviando todos los cálculos, tendríamos una imagen como la siguiente:

k	k-1	h(k)	ha(k)	I'	I
1	0	458	458	0	0
2	1	321	779 1,5922619048	2	
3	2	276	1055 2,9613095238	3	
4	3	230	1285 4,1021825397	4	
5	4	124	1409 4,7172619048	5	
6	5	57	1466	5	5

Por último asignamos a cada nivel el número de píxeles que le pertenecen tras ecualizar:

Nivel N°

Pixeles

1	458
2	0
3	321
4	276
5	230
6	181

**Para aplicar un cambio de brillo a una imagen tendremos que:**

{

~ Repetir pixeles.

= Aumentar o reducir el valor de cada píxel.

~ Promediar los valores para los componentes R,G,B de cada pixel.

}

Explicación: Repetir píxeles ocurriría si quisieramos ampliar la imagen ya que necesitas más píxeles que en la imagen original, promediar los valores R,G,B da como resultado una imagen en escala de grises, y la correcta sería aumentar el valor de cada pixel para modificar el brillo de la imagen. (Transparencia 5)

**Los tipos de detectores para puntos de esquina o "corners"**

{

- ~ están basados en un filtro "lineal" o "no lineal".
- = están basados en "edges" o en "niveles de gris".
- ~ no hay detectores para los puntos de esquina.

}

Explicación: Los puntos de esquina o "corners" son utilizados en un contorno de puntos de elevada curvatura o en uniones, puntos de intersección de varias aristas, normalmente con un significado geométrico. Los tipos de detectores, están basados:

- En "edges": a partir de la información de curvatura.
- En "niveles de gris": utilizan directamente el nivel de gris en una vecindad del píxel candidato para encontrar evidencia de que hay una esquina.

**Tras la aplicación de la convolución:**

{

- ~ El nuevo valor del píxel es el resultado de multiplicar la suma de los valores de la máscara por los valores de los píxeles.
- = El nuevo valor del píxel es el resultado de sumar la multiplicación de los valores de la máscara por los valores de los píxeles.
- ~ Ninguna de las anteriores es correcta.

}

**El ruido de una imagen son errores que pueden degradar la misma. Diremos que un método de procesamiento es robusto ante el ruido cuando:**

{

- = Cuando genera los mismos resultados con o sin ruido.
- ~ Cuando genera distintos resultados con o sin ruido.
- ~ Cuando los resultados son procesados pixel a pixel.

}

*Explicación: En la diapositiva 5 del tema se explica qué es el ruido, cuándo se produce y cuándo un método es robusto ante él, observando que la respuesta marcada como correcta, coincide con la explicación del tema.*

**Los errores que pueden degradar la imagen en su captura suelen ser:**

- {
- = Ruido
- ~ Cambio de brillo
- ~ Ampliar o reducir la imagen
- }

*Explicación: Aunque todas las soluciones pueden hacer que la imagen se degrade, en su captura solo afecta el ruido.*

**¿Qué filtro utiliza el proceso de detector de Canny para suavizar la imagen?:**

- {
- = Filtrado Gaussiano.
- ~ Filtro de media.
- ~ Filtro de mediana.
- }

*Explicación: La imagen se suavizada usando un filtro Gaussiano con una desviación estándar para reducir el ruido. Es un filtro lineal de media 0 y varianza  $\sigma$ .*

**¿Cuales son los criterios de la detección de Canny?:**

- {
- = Buena detección, buena localización y respuesta única.
- ~ Buena detección, buena localización y respuesta variada.
- ~ No importa el número de falsos positivos o negativos.
- }

*Explicación: Descartamos las opciones 2 y 3, puesto que la respuesta deben ser aristas*

*de un pixel de anchura, mientras que buscamos una mínima detección, por lo que debemos minimizar el número de falsos positivos y el de falsos negativos.*

**Para la construcción del detector de Nitzberg-Harris algunos sus pasos son:**

{

- = Calcular gradientes horizontal y vertical para cada uno de los píxeles de la vecindad
- ~ Aplicar promediado y registrar la matriz para algunos de los puntos.
- ~ Ninguna es correcta.

}

Explicación: En la diapositiva 33 del tema 11 podemos ver los unos de los pasos para la construcción del detector de Nitzberg-Harris es, calcular gradientes horizontal y vertical para cada uno de los píxeles de la vecindad, que es la respuesta correcta, ya que la segunda opción de aplicar promediado y registrar la matriz para algunos de los puntos, es incorrecta porque se debe aplicar promediado y registrar la matriz **para cada punto**.

**En percepción automática, en cuanto a los modelos de color:**

{

- ~ RGB no presenta problemas y siempre es mejor que HSV.
- = RGB presenta un problema importante, a la hora de segmentar por color, dos colores similares pueden aparecer lejos en el espacio de representación del modelo.
- ~ HSV (*red, green, blue*, del inglés), es un sistema de síntesis aditiva basado en los colores rojo, verde y azul.

}

*Explicación: HSV y RGB presentan problemas diferentes según el caso y no siempre es mejor RGB. HSV es un modelo de color basado en sus componentes del inglés Hue, Saturation, Value, es decir, Matiz, Saturación, Valor. La respuesta correcta se corresponde con un error típico del modelo RGB.*

**En cuanto al filtrado Gaussiano, indica la respuesta correcta**

{

- ~ Asigna al píxel central la media de todos los píxeles.
- ~ El valor mínimo aparece en el píxel central y aumenta hacia los extremos.
- = El valor máximo de la máscara de convolución aparece en el píxel central y disminuye hacia los extremos.

}

Explicación: El peso de los píxeles decrece con la distancia al centro, cuanto más alejado está un píxel, menos significativo es.

**A la hora de aplicar el detecto Canny:**

{

~ Nunca aplicamos un filtro de reducción de ruido, perderíamos detalles.

= Aplicamos el filtro Gaussiano, para reducir el ruido.

~ Aplicamos el filtro de Medias, para reducir el ruido.}

Explicación: Aplicamos una suavización sobre la imagen con la ayuda de un filtro gaussiano, de esta forma, reducimos el ruido y evitamos las aristas de tipo escalón.

**¿Cuál de las siguientes definiciones corresponde a un Edge (Arista)?**

{

~ Puntos de bajo contraste.

~ Puntos de baja derivada en valor absoluto.

= Puntos de alta derivada en valor absoluto.

}

Explicación: Un Edge o Arista es un punto de alta derivada en valor absoluto porque derivamos numéricamente una imagen (entendida como función bidimensional de variable real) para identificar las aristas. Sería incorrecto que fuera un punto de bajo contraste, ya que la intensidad de la imagen aumenta y disminuye bruscamente.

**En la ampliación de una imagen es cierto que:**

{

~ Ocupa menos memoria que la original.

~ Comprimimos la imagen.

= Se repiten pixeles.

}

Explicación:

La imagen resultado de la ampliación tiene más píxeles que la original y por lo tanto gasta más memoria. Hay que introducir nuevos píxeles en la imagen que deben ser interpolados a partir de los que ya tenemos. Como los nuevos píxeles son interpolados, es decir se calculan a partir de los que ya tenemos, la imagen ampliada no tiene más información que la original, y sin embargo gasta más memoria

**En cuanto a la ecualización del histograma, selecciona la respuesta correcta:**

{

- = Cuando tenemos una imagen con poco contraste podemos “expandir” el histograma.
- ~ Cuando tenemos una imagen con poco contraste podemos “reducir” el histograma.
- ~ Ninguna de las anteriores es correcta.

}

Explicación: Si tenemos una imagen con poco contraste no tiene sentido reducir el histograma porque no se conseguiría el resultado esperado, hay que expandirlo para conseguir el resultado correcto.

**En el detector de Canny, utilizando trade-off si aumentamos "sigma":**

{

- = reducimos ruido pero difuminamos los bordes y perdemos calidad en la localización.
- ~ aumentamos ruido pero acentuamos los bordes y ganamos calidad en la localización.
- ~ reducimos ruido pero difuminamos los bordes y ganamos calidad en la localización.

}

Explicación: según la diapositiva 30 del tema 11, en el detector de canny mediante el trade-off si aumentamos "sigma" reducimos ruido pero difuminamos los bordes y perdemos calidad en la localización.

**Hablando del detector de Canny, indica el orden correcto de trabajo de este sistema (ordena las afirmaciones):**

- A-Se define un punto de borde, como un punto cuyo peso es localmente máximo en la dirección del gradiente.
- B-La imagen se suavizada usando un filtro Gaussiano con una desviación estándar.
- C-Se realiza la unión de los pixeles, incorporando “candidatos débiles”.
- D-Los puntos de borde determinados anteriormente, originan crestas en la dirección

de crecimiento del gradiente de la imagen.

{

~ADCB.

~ACBD.

=BADC.

}

Explicación: El detector Canny es considerado, el más efectivo a la hora de detectar bordes, debido a su eficacia. Este detector, realiza su trabajo de la siguiente manera:

1. La imagen se suavizada usando un filtro Gaussiano con una desviación estándar, esto para reducir el ruido. El gradiente local, y la dirección del borde, son computadas en cada punto. El detector Sobel, puede ser usado para computar  $G_x$  y  $G_y$ .
2. Se define un punto de borde, como un punto cuyo peso es localmente máximo en la dirección del gradiente.
3. Los puntos de borde determinados en (2) originan crestas en la dirección de crecimiento del gradiente de la imagen. El algoritmo luego rastrea a lo largo de la cima de estas crestas, y lleva a cero los píxeles que no están en realidad sobre la cima de la cresta, originando una línea delgada en la salida, un proceso conocido como supresión no máxima. Los píxeles de crestas son luego comparados usando dos umbrales,  $T_1$  y  $T_2$  con  $T_1 < T_2$ . Los píxeles de crestas con valores mayores que  $T_2$  se dice que son “probables candidatos” para ser píxeles de borde. Los píxeles en crestas con valores entre  $T_1$  y  $T_2$  se dice que son “candidatos poco probables” para ser píxeles de borde.
4. Por último, el algoritmo realiza la unión de los píxeles, incorporando “candidatos débiles” que están 8-conectados a los píxeles “probables”.

**Cual es la respuesta correcta sobre el detector Nitzberg-Harris (matriz de momentos):**

{

= La matriz  $A(x, y)$  captura la estructura de la intensidad de la vecindad local.

~ Sean  $\lambda_1$  y  $\lambda_2$  los valores propios de la matriz  $A(x, y)$ . Los valores propios no forman una descripción rotacionalmente invariante.

~ Aplicar un promediado y registrar la matriz para un solo punto.

}

Explicación: La matriz captura la estructura de la intensidad de la vecindad local, la segunda es falsa porque los valores propios si forman una descripción rotacionalmente invariante, la tercera es incorrecta ya que se encarga de aplicar promediado y registrar la matriz para cada punto no para un solo punto.

**Aplicando un operador de Sobel de 3x3, obtenemos que la derivada de  $G_x$  es 0. ¿Habrá una transición en el Eje X?**

- {
- ~ Sí
- = No
- ~ No se puede determinar
- }

Explicación: En base a la siguiente fórmula:

$$\nabla I = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} f(x+1, y) - f(x, y) \\ f(x, y+1) - f(x, y) \end{bmatrix}$$

Si alguna de las derivadas es 0, es debido a que el elemento siguiente es igual al actual. Por tanto no se produce una transición.

**Si a una imagen le bajamos el contraste y le subimos el brillo. ¿Como será el nuevo histograma con respecto al de la imagen original?**

- {
- = Más comprimido y desplazado hacia la derecha.
- ~ Más comprimido y desplazado hacia la izquierda.
- ~ Más extendido y desplazado hacia la izquierda.
- }

Explicación: Si bajamos el contraste tendremos menos valores distintos de intensidad, por lo que el histograma se comprimirá, si aumentamos el brillo tendremos más valores más altos, por lo que el histograma se desplazara hacia la derecha.

**¿Qué detector es considerado el más efectivo a la hora de detectar bordes?**

- {
- = Canny
- ~ Nitzberg-Harris

~ Ninguno de los anteriores

}

Explicación: El detector Canny es considerado, el más efectivo a la hora de detectar bordes, debido a su eficacia. (Transparencia 23)

**En el proceso de captura de imagen mediante un sensor se tiene que:**

{

~ En los sistemas pin-hole se obtiene una clara ganancia de perspectiva.

= Se produce una discretización.

~ Las imágenes capturas se representan solo mediante unos.

}

*Explicación: Como se puede apreciar en la transparencia 3 del tema Introducción a la visión Artificial, al capturar una imagen mediante un sensor, se produce:*

- Sistemas CCD o CMOS (y otros)

- En los Sistema de pin-hole: se pierde la “perspectiva”

- Se produce una discretización

**Los supuestos de ruido gaussiano y arista de tipo "escalón" pertenecen a:**

{

= Detector de Canny.

~ Convolución.

~ Detector Nitzberg-Harris.

}

*Explicación: Es uno de los fundamentos del detector de Canny. Pag:21.*

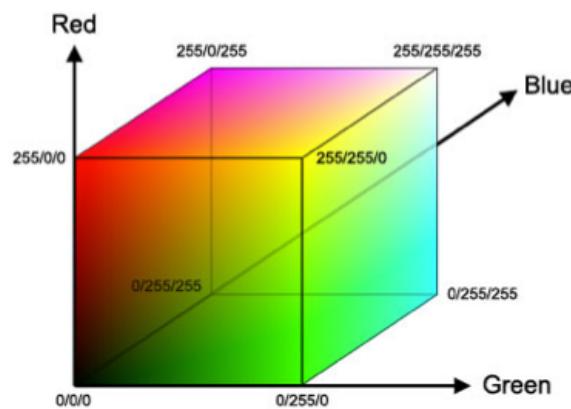
**¿Qué afirmación acerca de la percepción automática es falsa?**

{

- ~ Las imágenes de color RGB se representan con matrices de tres valores entre 0 y 255
  - ~ Cambiando el brillo de una imagen se aumenta o reduce el valor de cada píxel
  - = Si ampliamos o reducimos una imagen, el número de pixeles sigue siendo igual.
- }

*Explicación: La falsa es la ultima puesto que si se amplia una imagen se repiten pixeles y si se reduce se pierden pixeles.*

**El modelo de color RGB se representa con tres valores de 0 a 255 como se muestra en la imagen. ¿Por qué este modelo puede suponer un problema en el campo de los sistemas inteligentes?**



{

- = Porque existen colores muy similares a simple vista que tienen una representación RGB muy diferente (alejada), lo cual puede suponer un alto coste computacional.
  - ~ Porque existen colores muy similares a simple vista que tienen una representación RGB muy diferente (alejada), lo cual puede suponer un error de percepción para el sistema.
  - ~ Porque existen colores muy similares a simple vista que tienen una representación RGB muy parecida (próxima), lo cual puede suponer un alto coste computacional.
- }

Explicación: El modelo de color RGB puede suponer un problema en cuanto a coste computacional ya que existen colores que a simple vista son muy parecidos pero tienen una representación RGB muy diferente (alejada), cosa que no pasa con otros modelos de color. PD: en la imagen hay dos vértices con el mismo valor (0/255/255), creo que el de la izquierda de los dos debería ser 0/0/255. Corregidme si me equivoco.

**Si tenemos una fotografía con el siguiente histograma**



**diremos que la fotografía está:**

- {
- = Sobreexpuesta
- ~ Subexpuesta
- ~ Dentro de los valores de exposición normales
- }

Explicación: Este histograma indica sobreexposición ya que hay un gran número de pixeles luminosos, es decir, en la parte derecha del histograma.

**Respecto al algoritmo K-medias es cierto que:**

{

- = Es un método de agrupamiento heurístico con número de clases conocido (K).
- ~ Siempre encuentra la solución óptima.
- ~ Realiza un histograma de orientaciones alrededor de un punto.

}

*Explicación: El algoritmo de las K medias (o K-Means, J.B. MacQueen, 1967) es probablemente el algoritmo de agrupamiento más conocido. Es un método de agrupamiento heurístico con número de clases conocido (K). El algoritmo está basado en la minimización de la distancia interna (la suma de las distancias de los patrones asignados a un agrupamiento al centroide de dicho agrupamiento). De hecho, este algoritmo minimiza la suma de las distancias al cuadrado de cada patrón al centroide de su agrupamiento.*

**En el algoritmo para describir las características locales de las imágenes SIFT (escala invariante en función de transformación) publicado por David G. Lowe, indica cual sería el orden correcto:**

{

- ~Escala; Orientación; Localización del punto; Cálculo del descriptor.
- ~Escala; Localización del punto; Orientación; Cálculo del descriptor.
- =Localización del punto; Escala; Orientación; Cálculo del descriptor.

}

Explicación:

En el algoritmo SIFT el orden de los pasos a seguir en el tratamiento de una imagen es el siguiente: Localización del punto, escala, orientación y cálculo del descriptor.

**Según las características de K-medias, cual afirmación es falsa:**

{

- = En su modo probabilístico, se puede aplicar a un número reducido de distribuciones.
- ~ En su modo probabilístico, se pueden usar distintas medidas de distancia a distribución.
- ~ En su modo probabilístico, se puede utilizar con distancias como Mahalanobis y Kullback-Leiber.

}

Explicación: En su modo probabilístico, se pueden usar distintas medidas de distancia a distribución. como las distancias Mahalanobis y Kullback-Leiber. B y C son correctas.

Se puede aplicar a cualquier tipo de distribución (La A es falsa).

**¿Cual de las siguientes afirmaciones es correcta?:**

{

= SIFT tiene dos pasos: buscar los puntos característicos de una imagen y calcular sus descriptores.

~ SIFT no es multiescala

~ Los descriptores nos proporcionan información sobre los puntos característicos encontrados en una imagen, pero no permiten compararlos entre si.

}

Explicación: SIFT trabaja con la escala y la orientación. Tiene dos pasos: primero encuentra los puntos característicos de una imagen y después calcula sus descriptores para poder comparar unas características con otras.

### Acerca del mecanismo de la transformada de Hough puede afirmar que:

{

~ Cada punto de la imagen votará por aquella ecuación (combinación de parámetros, o celda) que cumpla. Daremos como salida las celdas sin suficiente soporte o evidencia, esto es, aquellas con un “número insuficiente de votos”.

= Los valores de los parámetros de la ecuación paramétrica definen únicamente a cada primitiva. Por cada parámetro tenemos una dimensión en el espacio paramétrico y cada dimensión se “discretiza” en celdas.

~ Ninguna de las anteriores.

}

Explicación: La primera opción no es correcta ya que la salida son las celdas CON suficiente soporte o evidencia. La segunda esta correcta, de acuerdo con que fue visto en clase.

### Cuál de los pasos siguientes NO pertenecen al algoritmo SIFT:

{

~ Encontrar la posición de los puntos;

~ Calcular el descriptor;

= Encontrar para cada punto el centroide más cercano.

}

Explicación: Encontrar para cada punto el centroide más cercano es un paso del algoritmo k-medias.

### Selecciona la respuesta correcta sobre los fundamentos de la Transformada de Hough:

{

= Espacio paramétrico. Los valores de los parámetros de la ecuación paramétrica definen únicamente a cada primitiva. Por cada parámetro tenemos una dimensión en el espacio paramétrico y cada dimensión se “discretiza” en celdas.

~ Espacio paramétrico. Los valores de los parámetros de la ecuación paramétrica definen ambiguamente a cada primitiva. Por cada parámetro tenemos dos dimensiones en el espacio paramétrico y cada dimensión se “discretiza” en celdas.

~ Espacio paramétrico. Los valores de los parámetros de la ecuación paramétrica definen a cada primitiva. Por cada parámetro no tenemos dimensiones en el espacio.

}

Explicación: Según los fundamentos de la Transformada de Hough la correcta es la primera, la segunda no puede ser porque no se definen de forma ambigua sino unívocamente y no tiene dos dimensiones por cada parámetro si no una, la tercera es incorrecta, ya que si se tienen dimensiones en el espacio por cada parámetro.

### **Selecciona la respuesta correcta acerca de la técnica de segmentación basada en regiones:**

{

~ no construye las regiones directamente.

= en general, trabaja mejor en imágenes con ruido.

~ busca cotas a partir de la distribución de propiedades de la imagen.

}

Explicación: las técnicas de segmentación basadas en regiones funcionan mejor cuando una imagen contiene ruido, ya que los bordes son difíciles de detectar.

### **Cuando aplicamos el algoritmo k-medias:**

{

~ Termina en k iteraciones.

= Termina cuando ya no hay cambios en las pertenencias o se han realizado el máximo de iteraciones fijadas como parámetro.

~ No podemos conocer cuando finaliza dicho algoritmo.

}

Explicación: Como hemos visto en prácticas el algoritmo de las k-medias se le pasa como parámetro el número máximo de iteraciones que puede ejecutarse para la búsqueda de la distribución más cercana para cada punto, pero puede llegar a estabilizarse antes de llegar a este número máximo de iteraciones.

### **¿Cuál de estas afirmaciones es incorrecta?**

{

~ El algoritmo de K-medias encuentra las medias de las distribuciones

= Cuando inicializamos el método de las K-medias no podemos redistribuir las medias de manera uniforme

~ El algoritmo de K-medias puede no encontrar la solución óptima

{

Explicación : Mediante el proceso de inicialización se puede

- Asignar las medias de manera aleatoria.

- Redistribuir las medias de manera uniforme

- Usar alguna heurística (por ejemplo, en secuencias de imágenes, usar las medias resultado anteriores)

Por lo que se deduce que la respuesta correcta es la segunda

### **Indica cuál de los siguientes NO es un paso del k-medias**

{

- ~ Calcular para cada punto, el clúster al que pertenece.

- = Asignaremos a cada clúster el centroide más cercano.

- ~ Ninguna de las anteriores.

}

Explicación: la asignación a cada clúster el centroide más cercano no corresponde al k-medias, es una cuantización posterior.

### **Cual de las siguientes afirmaciones NO es un punto critico en cuanto a la Transformada de Hough**

{

- ~ La complejidad espacial de la construcción del espacio de parámetros aumenta de forma factorial con el número de los mismos  $O(N!)$ .

- = Una mala inicialización puede llevar más tiempo.

- ~ Puede ser difícil de encontrar un buen ajuste de sensibilidad para evitar un elevado número de falsos positivos.

}

Explicación: En la diapositiva 7 del tema 12, se puede ver que la opción a) y c) constituyen puntos críticos en cuanto a la Transformada de Hough. Mientras que la opción b) se refiere a las características de las K-medias (diapositiva 31 del tema 12) y por eso es la respuesta correcta.

### **Respecto a la Transformada de Hough:**

{

- ~ Ineficiencia: la complejidad espacial de la construcción del espacio de parámetros disminuye de forma factorial con el número de los mismos:  $O(N!)$  Lo mismo sucede con la complejidad temporal del proceso de votación:  $O(N-1)!$ ;

- ~ Ineficiencia: la complejidad espacial de la construcción del espacio de parámetros aumenta de forma factorial con el número de los mismos:  $O(N!)$  Al contrario que con la complejidad temporal del proceso de votación:  $O(N-1)!$ ;

= Ajuste de sensibilidad: aunque es un factor de menor importancia, el ajuste del tamaño de celda y del número mínimo de votos es clave para evitar un elevado número de falsos positivos.

}

*Explicación:* Transparencia 7 del Tema 12 por definición.

### Cuáles son los puntos críticos de la transformada de hough?

{

= Ineficiencia, ajuste de sensibilidad y solamente es aplicable a primitivas sencillas;

~ eficiencia, ajuste de sensibilidad y solamente es aplicable a primitivas sencillas;

~Ineficiencia, ajuste de probabilidad y solamente es aplicable a primitivas sencillas;

} Explicación:

Puntos críticos:

- Ineficiencia: la complejidad espacial de la construcción del espacio de parámetros aumenta de forma factorial con el número de los mismos:  $O(N!)$  Lo mismo sucede con la complejidad temporal del proceso de votación:  $O(N-1)!$

- En consecuencia, el método exhaustivo solamente es aplicable a primitivas sencillas (caracterizadas por pocos parámetros) e incluso en estos casos se hace necesaria una discretización muy somera para satisfacer los requerimientos de memoria.

- Ajuste de sensibilidad: aunque es un factor de menor importancia, el ajuste del tamaño de celda y del número mínimo de votos (a menudo seguido de una supresión de no-máximos)

**En un problema de segmentación de imágenes y dadas las siguientes imágenes, ¿Qué**



**indicaría la variable "K"?:**

{

~ El valor HSV del punto central de la imagen

= El número de colores que tiene la imagen

~ Ninguna de las anteriores

}

*Explicación:* En un problema de segmentación de imágenes la variable "K" indica el número de colores que contiene la imagen. Por ejemplo para las imágenes dadas la variable "K"

sería 2, 4 y 8 respectivamente.

**Pregunta: Dentro del algoritmo Hough para el círculo, para cada (a,b,r) ... :**

{

~ Para cada pixel (x,y): r es la distancia manhattan entre a,b y x,y.

~ r es la diferencia entre a,b y la media de todos los x,y.

= Para cada pixel (x,y): r es la distancia euclídea entre a,b y x,y.

}

Explicación: r es la distancia euclídea entre a,b y x,y.

**Elige la opción correcta:**

{

~ El objetivo de la segmentación basada en regiones es encontrar regiones de la imagen heterogéneas según algún criterio.

~ SIFT calcula el receptor.

= Las medias de las distribuciones (clústeres) se pueden encontrar con el algoritmo de las K-Medias.

}

*Explicación: SIFT calcula el descriptor, con lo cual no es correcta. La segmentación basada en regiones debe encontrar regiones de la imagen homogéneas según algún criterio, también falsa. Las medias de las distribuciones (clústeres) se pueden encontrar con el algoritmo de las K-Medias es la opción correcta.*

**En relación con la percepción automática y en concreto con los puntos críticos de la transformada de Hough, señala la respuesta correcta:**

{

~ Ineficiencia: las complejidades espaciales y temporales son  $O(N!)$  y  $O(N^2 - 1)$  respectivamente.

~ Ineficiencia: las complejidades temporales y espaciales son  $O(N!)$  y  $O(N-1)!$  respectivamente.

= Ineficiencia: las complejidades espaciales y temporales son  $O(N!)$  y  $O(N-1)!$  respectivamente.

}

*Explicación:*

*Los mecanismos de la transformada de Hough son:*

1.- Espacio paramétrico. Los valores de los parámetros de la ecuación paramétrica definen unívocamente a cada primitiva. Por cada parámetro tenemos una dimensión en el espacio paramétrico y cada

dimensión se “discretiza” en celdas.

2.- Proceso de “votación”. Cada punto de la imagen votará por aquella ecuación (combinación de parámetros, o celda) que cumpla.

Daremos como salida las celdas con suficiente soporte o evidencia, esto es, aquellas con un “número suficiente de votos”.

*Uno de los puntos críticos es la ineficiencia ya que:*

La complejidad espacial de la construcción del espacio de parámetros aumenta de forma factorial con el número de los mismos:  $O(N!)$  Lo mismo sucede con la complejidad temporal del proceso de votación:  $O(N-1)!$

*Por lo tanto la respuesta correcta es la tercera*

**En cuanto a la segmentación basada en regiones de una imagen, indica la respuesta correcta**

{

= En el método de Crecimiento de regiones podemos empezar por cualquier píxel de la imagen pero es mejor lanzar varios puntos de partida "semillas"

~ Según el criterio de Partición de regiones aplicamos un detector de aristas para buscar los píxeles iniciales.

~ En el crecimiento de regiones utilizamos un test de similaridad en el que si un píxel no lo cumple se descarta y no forma parte de ninguna región.

}

*Explicación: Estas afirmaciones hacen referencia a las páginas 34, 35 y 36 del tema 12.*

*La primera afirmación, la correcta, podemos ver que en la página 35 nos dice que el método crecimiento de regiones se puede empezar por cualquier píxel de la imagen, pero que existe la variante con "semillas" que lanza varios puntos de partida mejorándolo.*

*En cuanto a la segunda, se refiere al crecimiento de regiones con semillas, como podemos ver en la página 36 que trata de la partición de regiones no se hace referencia a los detectores de aristas. Además no se inicia en ningún píxel, se parte de una única región (toda la imagen)*

*La tercera, es sobre la página 34, es cierto que existe un test de similaridad para el crecimiento de regiones, pero si el pixel al que se lo aplicamos no lo cumple lo que haremos es empezar a calcular una nueva región con ese píxel.*

**Selecciona la respuesta correcta sobre la segmentación basada en regiones:**

{

= El objetivo es encontrar regiones de la imagen homogéneas según algún criterio.

~ En el crecimiento de regiones, se asume que una región es un conjunto de píxeles sin ningún tipo de conexión.

~ El objetivo es encontrar regiones de la imagen heterogéneas según algún criterio.

}

Explicación: La segmentación basada en regiones intenta encontrar regiones de la imagen homogéneas no heterogéneas, y en el crecimiento de regiones asume que una región es un conjunto de píxeles conectados.

### **En cuanto a los puntos críticos de la transformada de Hough**

{

- ~ La complejidad temporal del proceso de votación aumenta de forma factorial  $O(N!)$
- ~ El ajuste del tamaño de celda y del número mínimos de votos es clave para evitar un elevado número de falsos negativos
- = El método exhaustivo solamente es aplicable a primitivas sencillas, haciéndose necesaria una discretización ligera para satisfacer los requerimientos de memoria.

}

Explicación: La solución es la c), la complejidad temporal es  $O(N-1)!$  y la espacial  $O(N!)$ . La b) no es porque ese ajuste es para evitar falsos positivos. Dada la gran complejidad que supone el método exhaustivo, sólo es aplicable para primitivas con pocos parámetros. Aun en esos casos se debe disminuir un poco la "carga" discretizando ligeramente.

### **Determinar la definición correcta sobre la segmentación de imágenes:**

{

- ~ Es el proceso de extraer zonas de la imagen con el mismo tamaño para clasificarlas automáticamente.
- = Es el proceso de extraer zonas de la imagen con el mismo color/nivel de gris/textura para identificarlas automáticamente.
- ~ Ninguna de las anteriores es correcta.

}

Explicación: La primera opción es incorrecta por que ni se extraen zonas por tamaño ni se clasifican, la segunda es correcta por que coincide exactamente con la definición de la transparencia 24, y por esto mismo la tercera es imposible.

### **¿Cómo se consigue la localización en la multiescala?:**

{

- ~ Con una suma de gaussianas (SoG).
- = Con una diferencia de gaussianas (DoG).
- ~ Con una división de gaussianas (DoG).

{

Explicación: Por definición, la localización en la multiescala se consigue con una diferencia de gaussianas (DoG).

Tema 12 - Pág. 14

**En la segmentación basada en regiones la manera en la que empezamos con regiones pequeñas y las hacemos crecer o bien las mezclamos, usando un criterio de similaridad se llama:**

{

- = Crecimiento de regiones.
- ~ Partición de regiones.
- ~ Ninguna de las anteriores.

}

Explicación: Transparencia 32 del Tema 12, la definición pertenece a Crecimiento de regiones, ya que Partición de regiones empieza con regiones grandes y las vamos dividiendo usando un criterio de homogeneidad.

**¿Qué devuelve el algoritmo SIFT?:**

{

- ~ Un booleano.
- = Las características encontradas en la imagen.
- ~ Ninguna de las otras respuestas es correcta.

}

Explicación: Tal y como aparece en la página 11 del tema 12, al emplear este algoritmo recibimos como respuestas las características de la imagen.

**El algoritmo de las K-Media se inicializa:**

{

- = Asignando las medias de manera aleatoria.
- ~ Redistribuyendo las medias de forma desigual.
- ~ Ninguna de las anteriores es correcta.

}

*Explicación: El Algoritmo encuentra las medias de las distribuciones (clústeres) y es necesario conocer el número de distribuciones (clústeres) distintas existentes. Para ello utiliza una heurística que debe ser calculada en la fase de inicialización*

**Indica qué secuencia sigue el algoritmo SIFT para encontrar las características de una imagen:**

{

- = Encontrar la posición de los puntos, calcular el descriptor.
- ~ Calcular el descriptor, encontrar la posición de los puntos.
- ~ Encontrar las zonas más saturadas de la imagen, determinar los puntos en esas zonas, calcular el descriptor.

}

*Explicación: La secuencia que sigue el algoritmo SIFT es: Encontrar la posición de los puntos, calcular el descriptor.*

**En la segmentación basada en regiones existen dos maneras de hacerlo, ¿cuáles son?**

{

- ~ Partición de regiones y filtración de regiones.
- = Crecimiento de regiones y partición de regiones.
- ~ Ninguna de las anteriores.

}

*Explicación: Las 2 maneras que existen son las de crecimiento y partición ya que, para el crecimiento se empieza desde regiones pequeñas y van creciendo; y la partición de regiones se empieza con regiones grandes y van decreciendo.*

**En segmentación basada en regiones, en partición de regiones (*region splitting*):**

{

- = Empezamos con regiones grandes y las vamos dividiendo usando un criterio de homogeneidad.
- ~ Empezamos con regiones pequeñas y las vamos dividiendo usando un criterio de homogeneidad.

- ~ Empezamos con regiones grandes y las vamos dividiendo usando un criterio de heterogeneidad. }

Explicación: Transparencia 32 del Tema 12, la definición correcta de partición de regiones es empieza con regiones GRANDES y las vamos divide usando un criterio de HOMOGENEIDAD.

### **En la transformada de Hough, la motivación es:**

{

- ~ Desarrollar técnicas que permitan identificar formas geométricas complejas en una imagen.
- = Desarrollar técnicas que permitan identificar primitivas geométricas sencillas en una imagen.
- ~ Ninguna de las anteriores.

}

Explicación: Como podemos leer en la diapositiva 3 del tema 12, la motivación de la transformada de Hough es desarrollar técnicas que permitan identificar primitivas geométricas sencillas (líneas, círculos, cuadrados....) en la imagen (p.e. tras binarizar el resultado de Canny).

### **Para hacer uso del algoritmo de las K-medias:**

{

- ~ Es necesario conocer varias de las distribuciones distintas existentes.
- = Es necesario conocer el número de distribuciones distintas existentes.
- ~ No es necesario conocer el número de distribuciones distintas existentes.

}

Explicación: en la transparencia 25 del tema 12; "En el algoritmo de las K-medias es necesario conocer el número de distribuciones (clústeres) distintas existentes".

### **En lo referente a la localización, ¿Cuáles son los puntos relevantes?**

{

- ~ Aquellos que son inestables.

= Los máximos o mínimos.

~ No hay puntos relevantes.

}

*Explicación: Los puntos relevantes son los máximos o mínimos respecto de su vecindad, de la escala inferior o superior (los encontramos entre escalas) y no son inestables. Estos marcan la diferencia entre el plano superior y el plano inferior.*

### **En cuanto a la partición de regiones:**

{

= Conlleva una complejidad mayor, al tener que manejar alguna estructura de datos adicional.

~ Conlleva una complejidad menor, al tener que manejar alguna estructura de datos adicional.

~ Conlleva una complejidad mayor, al tener que manejar un criterio de homogeneidad.

}

Explicación: En la página 36 del tema pone en la diapositiva sobre Partición de regiones se indica que conlleva una complejidad mayor, al tener que manejar alguna estructura de datos adicional.

### **En cuanto al crecimiento de regiones mediante semillas:**

{

= Se aplica un detector de aristas. Los puntos cuyo valor de magnitud de gradiente estén próximos a cero, serán “valles”, es decir, puntos dentro de regiones. Usaremos estos puntos como “semillas”.

~ Lo mejor es coger un único punto de partida o semilla, como puede ser una esquina, y a partir de este punto ir creando todo.

~ Ninguna de las anteriores es correcta.

}

Explicación: En la página 35 del tema. Se nos explica exactamente esta descripción. Si cogiésemos un único punto de partida, el coste de crear todas las semillas y sus regiones es demasiado alto. Puesto que la opción 1 es correcta la 3 no puede ser.

**La segmentación de imágenes es el proceso de extraer zonas de la imagen para poder identificarlas automáticamente ¿En base a que criterios?**

{

- = Mismo color/nivel de gris/textura.
- ~ Mismo color/nivel de gris/orientación.
- ~ Mismo color/nivel de gris/contraste.

}

Explicación: La segmentación de imágenes es el proceso de extraer zonas de la imagen con el mismo color/nivel de gris/textura para identificarlas automáticamente.

Tema 12 - Transparencia 24

**¿Que pega puede tener la partición de regiones respecto a otros?**

{

- ~ Su inexactitud respecto a otros
- = Complejidad mayor al tener que manejar alguna estructura de datos adicional
- ~ Ninguna de las anteriores es correcta

}

Explicación: La partición de regiones tiene una pega respecto a otros, y es su complejidad mayor al tener que manejar alguna estructura de datos adicional.

**Una diferencia entre la transformada de Hough y SIFT:**

{

- = SIFT trabaja de forma correcta sobre imágenes con diferente iluminación, ángulo de captura, etc.
- ~ Hough es el más adecuado para tratar imágenes con distinta iluminación, ángulo de captura, etc.
- ~ Hough hace uso de descriptores, pero SIFT no.

}

Explicación: SIFT, al trabajar con descriptores y no con puntos, hace uso de la información

que existe no solo en un punto, si no alrededor de este. Esto permite obtener las mismas características sobre una misma imagen tomada en diferente ángulo, iluminación, etc.

**Selecciona la opción que no se adapte (incorrecta) a la hora de la resolución de problemas de cambios de iluminación:**

{

~ Uso de descriptores que contengan información necesaria para identificar los distintos puntos en una imagen.

= Emplear la transformada de Hough para determinar el reconocimiento de patrones geométricos inherentes a la imagen es suficiente para resolver problemas de iluminación.

~ El algoritmo de SIFT

}

**En un problema de segmentación de imágenes y dadas las siguientes imágenes, ¿Qué indicaría la variable "K"?:**

{



~ El valor HSV del punto central de la imagen

= El número de colores que tiene la imagen

~ Ninguna de las anteriores

}

Explicación: En un problema de segmentación de imágenes la variable "K" indica el número de colores que contiene la imagen. Por ejemplo para las imágenes dadas la variable "K" sería 2, 4 y 8 respectivamente.

**Hablando del algoritmo de las K-medias, elige la opción falsa**

{

~El algoritmo de las k-medias se puede emplear para clasificar los descriptores devueltos por SIFT en clúster.

~Puesto que la segmentación de imágenes se basa la discontinuidad del nivel de gris en las imágenes por una parte. Por otra parte en la continuidad del nivel de gris, el algoritmo de las k-medias se puede emplear para la segmentación de imágenes usando el nivel de gris esta vez.

=La transformada de Hough usa coordenadas polares para la identificación de primitivas geométricas sencillas de imagen. Por lo que el algoritmo de las K-medias se puede usar también.

}

Explicación:

opción 1 no responde a la pregunta: En la práctica 2 de sistemas Inteligentes curso 2012/2013, utilizamos las K-medias para agrupar los descriptores en clúster.

Opción 2 tampoco responde a la pregunta. En efecto, la segmentación simplifica y/o cambia la representación de una imagen en otra más significativa y fácil de analizar. Para ello usa los niveles de gris (fuente Wikipedia). También, se puede consultar la pagina 30 de este tema.

Opción 3 responde a la pregunta. En efecto, la transformada de Hough identifica las primitivas geométricas, no mueve los pixeles. Pagina 3 del tema.

**Con el algoritmo K-means conseguimos asignar cada punto de una imagen a un clúster basándonos en las características de dicho punto pero, ¿Qué representan los clústeres?**

{

~ Una región del espacio que hemos de seguir particionando

= Son las medias de las distribuciones a partir de las cuales podemos clasificar una imagen con un conocimiento previo

~ Un segmento de la imagen característico

}

Explicación: K-means es un algoritmo de aprendizaje que bien puede ser usado como clasificador (es lo que estamos haciendo, de hecho, en la práctica) con lo que solo la segunda opción es correcta

**¿Qué afirmación acerca del algoritmo de k-medias es verdadera?**

{

= Se buscan k puntos (medias) y es necesario indicar explícitamente el valor de k

- ~ El algoritmo solo se puede inicializar de una forma en concreto.
  - ~ Nunca encuentra la solución más óptima
- }

Explicación: El algoritmo puede o no encontrar la solución óptima y puede inicializarse de diferentes formas como de forma aleatoria o heurística, por lo que la 2 opción y la 3 son falsas. La primera es la correcta.

**Para la transformada de Hough, es cierto que:**

- {
- ~ Permite identificar texturas
  - . = Cada punto de la imagen participa en un proceso de votación.
  - ~ Se utiliza por su bajo coste computacional. } Respuesta: La correcta es la segunda: cada punto de la imagen participa en un proceso de votación.

Explicación: La primera es falsa porque permite identificar primitivas geométricas sencillas como líneas, círculos o cuadrados. La tercera es falsa porque su coste es factorial.

**¿Que problema plantea utilizar SIFT para el reconocimiento de caras? {**

~Es un algoritmo que no consigue reconocer caras.

~El algoritmo SIFT no plantea problemas, puesto que es el que se utiliza. De hecho es el más rápido.

=No nos da una respuesta suficientemente rápida para poder aplicarlo a cámaras, y otros captadores de imagen.}

Explicación: En clase se nos ha explicado que el SIFT aunque si que consigue reconocer caras, no es suficientemente rápido para poder ser utilizado.

**En las Características de región en Rectángulos, nos encontramos ante una determinada región cuyo****valor ( $\sum(\text{píxeles área negra}) - \sum(\text{píxeles área blanca}) = 0$ , ¿qué podemos determinar? {**

~ Es una región Heterogénea.

= Es una región Homogénea.

~ Es una región indeterminada.

}

Explicación: Nos encontraríamos ante una región cuyos valores de píxeles negros y blancos se neutralizarían, dando lugar a una región Homogénea.

**Sobre "The Attentional Cascade", para definir este mecanismo, hay que determinar: {**

~ El número de niveles, el número de características en cada nivel y elegir el clasificador con mayor error.

~ El número de niveles, normalizar los pesos y el umbral de cada nivel.

= El número de niveles, el número de características en cada nivel y el umbral de cada clasificador.

}

Explicación: Para definir este mecanismo hay que determinar:

El número de niveles.

El número de características de cada nivel.

El umbral de cada clasificador.

Las dos anteriores no pueden ser, ya que la primera nos habla de elegir el clasificador con mayor error y es incorrecto, la segunda también ya que no hay que normalizar los pesos ni determinar el umbral de cada nivel.

**Sobre el Algoritmo de Viola&Jones de reconocimiento de caras, selecciona la opción correcta: {**

~ El método tiene un porcentaje medio de falsos positivos es decir su credibilidad es media

~ No le afecta el cambio de luminosidad

= Sólo sirve para caras frontales o con poco giro

}

Explicación: Como podemos ver en la diapositiva 10 del tema 13. El Algoritmo de Viola&Jones de

reconocimiento de caras tiene las siguientes características:

- El objetivo es detectar caras, no reconocerlas
- Otro objetivo es que sea muy rápido: este método tarda pocos milisegundos en procesar una imagen
- El método tiene pocos falsos positivos y un alto porcentaje de detección correcta
- Sólo sirve para caras frontales o con poco giro
- Le afecta el cambio de luminosidad

Por estas características podemos decir que la correcta es que sólo sirve para caras frontales o con poco giro, ya que el cambio de luminosidad si que le afecta y el método tiene pocos falsos positivos y un alto porcentaje de detección correcta, no como dice la primera respuesta que dice que tiene un porcentaje medio de falsos positivos.

### **Sobre el reconocimiento de características el emparejamiento se calcula: {**

- ~ Eliminando los pixeles idénticos.
- = Utilizando la distancia euclídea del descriptor.
- ~ Añadiendo pixeles del mismo color.

}

Explicación: en la Diapositiva 6 del tema 13 vemos que el emparejamiento se puede hacer calculando la distancia euclídea del descriptor.

Además, para cada característica del modelo:

- Encontramos la característica de la escena cuya distancia euclídea esté por debajo de un cierto umbral.
- Ahora tenemos una correspondencia entre los descriptores del modelo y de la escena.

### **Acerca del attentional cascade. ¿Cuál de las siguientes características no hay que determinar?{**

- ~ El número de niveles
- = El umbral de cada nivel
- ~ Número de características en cada nivel

}

Explicación: Como podemos observar en la transparencia 25, hay que determinar el umbral de cada clasificador y no de cada nivel.

### **En el reconocimiento de objetos, ¿qué ventaja tiene la imagen integral?: {**

- ~ Como lo que queremos calcular es la suma de los píxeles dentro de un rectángulo, la imagen integral permite hacer este cálculo con una operación básica.
- ~ Como lo que queremos calcular es la suma de los píxeles dentro de un rectángulo, la imagen integral permite hacer este cálculo con dos operaciones básicas.
- = Como lo que queremos calcular es la suma de los píxeles dentro de un rectángulo, la imagen

integral permite hacer este cálculo con tres operaciones básicas.

}

Explicación: Por definición, en el reconocimiento de objetos, como lo que queremos calcular es la suma de los píxeles dentro de un rectángulo, la imagen integral permite hacer este cálculo con tres operaciones básicas.

Tema 13 - Pág. 16

**En cuanto a attentional cascade, sabiendo que el esquema usado crea un conjunto de clasificadores por niveles:** {

~ En el primer nivel, no se entrena ningún clasificador.

= Los píxeles que sobreviven a un nivel pasan al siguiente, que tendrá otro clasificador con más características, es decir, más especializado.

~ Los clasificadores que sobrevivan a un nivel pasan al siguiente nivel.

}

*Explicación: como hemos comprobado en el tema En el primer nivel, se entrena uno con pocas características, rápido, que permita eliminar la mayoría de no caras y los que sobrevivan pasan al siguiente nivel, que tendrá otro clasificador con más características, pero más especializado.*

Tema13- Diapositiva 24

**Respecto al concepto de "Imagen Integral", ¿cuál de las siguientes afirmaciones es CORRECTA?:**

{

~ Cada punto en la imagen integral es la suma de todos los píxeles de la imagen original a la derecha y arriba de ese punto.

~ Cada punto en la imagen original es la suma de todos los píxeles de la imagen integral encima y a la izquierda de ese punto.

= Cada punto en la imagen integral es la suma de todos los píxeles de la imagen original a la izquierda y arriba respecto a ese punto.

}

*Explicación: Como viene explicado en la diapositiva 14 del Tema 13, la imagen integral consta de una colección de puntos en la que, cada uno de ellos está compuesto por la suma de píxeles situados arriba y a la izquierda de dicho punto.*

**Respecto al método de Viola & Jones para reconocimiento de caras, indica la opción correcta**

{

~ El objetivo es detectar caras y reconocerlas.

= Sólo sirve para caras frontales o con poco giro.

~ Ninguna de las anteriores.

}

Explicación: El objetivo es detectar caras, no reconocerlas, por lo cual sólo habría una respuesta correcta, la segunda, ya que sólo sirve para caras frontales o con poco giro para el reconocimiento de caras.

Transparencia 10.

**En el algoritmo de Viola&Jones: Robust Real-Time Object Detection, ¿Cómo se mejoró la eficiencia de AdaBoost a la hora de reconocer caras en fotografías?: {**

- ~ Mediante la eliminación aleatoria de zonas de la fotografía.
- = Aplicación de los clasificadores por niveles. Aumentando el número de clasificadores en cada iteración.
- ~ Aplicación de los clasificadores por niveles. Disminuyendo el número de clasificadores en cada iteración.

{}

Explicación: Una forma de aumentar considerablemente el algoritmo de proceso de la imagen, se aplican en una primera vuelta únicamente dos clasificadores (relacionados con nariz y ojos). A continuación, a las zonas donde es posible que exista una cara (aproximadamente, de forma empírica, se trata de un 50% de la foto) se les vuelve a aplicar el algoritmo, pero con un número mayor de características.

De esta forma, obtenemos un reconocimiento con un gran valor de acierto, y un menor coste de procesamiento.

**¿Cuál de las siguientes opciones NO dificulta el reconocimiento de objetos? {**

- = Extensión del fichero de imagen

- ~ Cambio de iluminación

- ~ Escala

{}

Explicación: La extensión del fichero de imagen no es un factor que dificulte el reconocimiento de objetos, solo indica el formato del fichero. No debemos confundir la extensión del fichero con la compresión aplicada en la imagen ya que esta sí dificulta el reconocimiento de objetos al suprimir parte de la información original de una imagen.

**¿Cuál puede ser un objetivo del algoritmo Viola&Jones?{**

- ~ El objetivo es reconocer caras, no detectarlas

- = Que sea muy rápido

- ~ Ninguna de las anteriores es cierta

{}

Explicación: El algoritmo de Viola&Jones (algoritmo para reconocimiento de caras) su principal objetivo es el de detectar caras, no reconocerlas. Además como cualquier otro algoritmo se pretende que sea lo más rápido posible.

**El cálculo de la imagen Integral se puede realizar: {**

- ~ en tres pasadas.

- ~ en dos pasadas.

- = en una sola pasada.

{}

**Explicación:** Se puede calcular en un solo paso:

$$s(x, y) = s(x, y-1) + i(x, y)$$

$$ii(x, y) = ii(x-1, y) + s(x, y)$$

donde  $s(x,y)$  es la suma acumulada de la columna. Hay que tener en cuenta las primeras filas y columnas.

**¿De qué complejidad es el cálculo de la imagen integral completa? {**

~ Cuadrática

= Lineal

~ Constante

}

Explicación: El cálculo de la imagen integral es lineal puesto que si se tienen las acumulaciones obtenidas en el píxel superior y en el izquierdo, únicamente se realiza una suma más para la obtención del píxel actual. Por lo anterior, el tiempo empleado en el cálculo de la imagen integral completa depende del número de píxeles que tenga.

**¿Qué objetivo tiene el reconocimiento de caras en el algoritmo Viola&Jones?:{**

~ Reconocer una cara concreta

= Detectar una cara

~ Ninguna de las anteriores

}

Explicación: Según la transparencia nº 10 del tema, el objetivo del reconocimiento en este algoritmo es detectar una cara, no reconocerlas, con el objetivo de que sea un método muy rápido.

**Respecto al cálculo de la imagen integral, indica la respuesta correcta: {**

~ Es el método encargado de seleccionar las características y de conseguir un clasificador resultante de la combinación de clasificadores débiles.

~ Surge de la necesidad de obtener rapidez de cálculo y obtiene cada valor realizando:  $\Sigma(\text{píxeles área negra}) - \Sigma(\text{píxeles área blanca})$ .

= Cada punto en la imagen integral es la suma de todos los píxeles de la imagen original a la izquierda y arriba de ese punto.

}

Explicación: La primera opción es incorrecta debido a que está definiendo el funcionamiento de Adaboost. En la segunda opción, en cambio, existe una parte que la hace correcta, pero la obtención del resultado se corresponde con el método de extracción de características empleando rectángulos. La tercera opción se puede ver en la página 13 del tema, y es la idea que sustenta la eficacia del método en cuestión; y por tanto correcta.

**El algoritmo de Viola&Jones...{**

- ~ Su objetivo principal es la de detectar caras y reconocerlas.
- ~ Es muy rápido, solamente tarda pocos segundos en procesar una imagen.
- = No puede reconocer las caras.

{}

Explicación, Como podemos ver en la página 10 del tema 13, la única opción correcta es la tercera ya que este algoritmo no puede reconocer las caras, sólo nos sirve para poder detectarlas. Si quisiéramos reconocerlas deberíamos utilizar otras técnicas de reconocimiento y tener una base de datos para poder asociarlas.

**En el algoritmo de Viola&Jones visto en clase, indica cuales son sus objetivos principales:{**

- ~Reducir falsos positivos. Que no afecte la luminosidad.
- ~Reconocer caras frontales y giradas. Reducir falsos positivos.
- =Que sea muy rápido. Detectar caras, no reconocerlas.

{}

Explicación:

En el algoritmo de Viola&Jones, el objetivo es detectar caras, no reconocerlas otro objetivo es que sea muy rápido: este método tarda pocos milisegundos en procesar una imagen. Además el método tiene pocos falsos positivos y un alto porcentaje de detección correcta, sólo sirve para caras frontales o con poco giro y le afecta el cambio de luminosidad.

**Sobre "The Attentional Cascade":**

{

- ~ De media, se necesita más procesamiento para eliminar no caras.
- = De media, se necesita menos procesamiento para eliminar no caras.
- ~ Ninguna de las dos es correcta.

}

Explicación: Tal y como figura en la página 24 del tema 13, "De media, se necesita menos procesamiento para eliminar no caras".

**¿Qué afirmación acerca de la percepción automática es verdadera? {**

- ~ El reconocimiento de objetos es un proceso sencillo
- ~ El reconocimiento de objetos es un proceso difícil pero no le afectan los cambios de iluminación
- = El reconocimiento de objetos es un proceso difícil debido entre otras cosas a la presencia de otros objetos no modelados.

}

Explicación: La verdadera es la última puesto que el reconocimiento de objetos es un proceso

difícil debido a algunos factores como: presencia de otros objetos no modelados, cambio de iluminación, cambio de punto de vista del objeto, ocultación y escala.

### En cuanto al método de Viola&Jones para detección de caras {

- ~ Tiene como objetivo asociar una cara a una persona
  - ~ Sirve para cualquier tipo de cara, sea cual sea su orientación o giro
  - = Es un método muy rápido.
- }

Explicación: La solución es la c), este método tarda pocos milisegundos en procesar una imagen. La a) no es porque el método trata de encontrar las caras de la imagen, no de reconocerlas. La b) no es porque, debido a que se basa sobre todo en la relación nariz-ojos, sólo sirve para caras frontales o con poco giro.

### En lo referente al Algoritmo de Viola & Jones. Selecciona la respuesta correcta ...{

- ~ No se ve afectado por los cambios de luminosidad
  - ~ Detecta fácilmente caras de perfil
  - = El objetivo es detectar caras, no reconocerlas.
- }

Explicación: Sólo sirve para caras frontales o con poco giro y tiene problemas con los cambios de luminosidad, por lo que la respuesta correcta es la última.

Encontramos las características del Algoritmo en la página 10 del T13.

### ¿Qué algoritmo se utiliza para calcular la selección de características de una imagen integral?

{

- ~ Bagging
  - = Adaboost.
  - ~ Ninguno en concreto, se utilizará la técnica o algoritmo que el programador crea conveniente.
- }

Explicación: Para seleccionar las características en una imagen integral debemos utilizar el algoritmo Adaboost para seleccionar solo aquellas que sean relevantes, ya que por ejemplo en el caso de una ventana de 24x24 se forman 180000 combinaciones de tamaño y orientación.

### Sobre el algoritmo Viola&Jones podemos decir que:{

- ~ El objetivo es detectar y reconocer caras.
  - = El método tiene pocos falsos positivos y un alto porcentaje de detección correcta.
  - ~ No le afecta el cambio de luminosidad.
- }

Explicación: La primera opción no es correcta ya que el objetivo es detectar caras, no

reconocerlas.

*La tercera es incorrecta porque le afecta el cambio de luminosidad.*

*La segunda esta correcta, de acuerdo con que fue visto en clase.*

B se tiene que quitar la suma de los píxeles dentro del rectángulo A (II(1)).

**En relación con la percepción automática, el reconocimiento de objetos consiste en, dado algún conocimiento (forma, apariencia, etc.) sobre uno o varios objetos y una imagen, encontrar qué objetos están en la imagen y dónde. El reconocimiento es un proceso difícil debido a (señala la respuesta correcta):{**

= Presencia de otros objetos no modelados, cambio de iluminación, cambio de punto de vista del objeto, oclusión, escala.

~ Presencia de otros objetos no modelados, cambio de punto de vista del objeto, oclusión, escala. El cambio de iluminación no es un factor importante.

~ Presencia de otros objetos no modelados, cambio de iluminación, cambio de punto de vista del objeto, oclusión. El cambio de escala no es un factor importante.

}

Explicación: en la diapositiva 3 del Tema 13 de percepción automática, se menciona que el reconocimiento es una tarea complicada debido a los siguientes factores; a la presencia de otros objetos no modelados, al cambio de iluminación, al cambio de punto de vista del objeto, la oclusión y la escala.

**Uno de los factores por lo que el reconocimiento de objetos es un proceso difícil es :{**

~ La presencia de otros objetos modelados.

= Al cambio de punto de vista del objeto.

~ Ninguna de las anteriores es correcta.

}

Explicación: según la transparencia 3 del tema 13, "El reconocimiento de objetos es un proceso difícil debido a la presencia de otros objetos no modelados, al cambio de iluminación, al cambio de punto de vista del objeto, la oclusión y la escala."

**En el reconocimiento con características se hace uso de un modelo del que se extraen sus características SIFT por las que dicho modelo se verá representado. Si ahora tenemos una nueva imagen ¿con qué método podemos encontrar las correspondencias entre las características del modelo y las de la nueva imagen? {**

~ Extrayendo las características SIFT de la nueva imagen y calculando la distancia euclídea del descriptor. Para UNA característica del modelo: cuando se encuentra una característica de la nueva imagen cuya distancia euclídea esté por DEBAJO de un cierto umbral lo que tendremos será una correspondencia entre los descriptores del modelo y los de la nueva escena.

~ Extrayendo las características SIFT de la nueva imagen y calculando la distancia euclídea del descriptor. Para CADA característica del modelo: cuando se encuentra una característica de la

nueva imagen cuya distancia euclídea esté por ENCIMA de un cierto umbral lo que tendremos será una correspondencia entre los descriptores del modelo y los de la nueva escena.

= Extrayendo las características SIFT de la nueva imagen y calculando la distancia euclídea del descriptor. Para CADA característica del modelo: cuando se encuentra una característica de la nueva imagen cuya distancia euclídea esté por DEBAJO de un cierto umbral lo que tendremos será una correspondencia entre los descriptores del modelo y los de la nueva escena.

}

Explicación: En el reconocimiento con características, el emparejamiento de las mismas se puede hacer calculando la distancia euclídea del descriptor. Esto se hace para cada una de las características del modelo y solo cuando la distancia euclídea de una característica de la escena esté por debajo de un cierto umbral tendremos una correspondencia entre los descriptores del modelo y de la escena.

### **¿Cuál es correcta de estas afirmaciones?**

{

~ El reconocimiento de objetos se hace más difícil por varios motivos uno de ellos es por la presencia de otros objetos modelados

~ El cambio de Iluminación no afecta el reconocimiento de objetos

= La oclusión es un problema que dificulta el reconocimiento de objetos.

}

Explicación: La primera es incorrecta dado que lo que hace más difícil el reconocimiento de objetos son los otros objetos NO modelados.

Los cambios de luz siempre afectan para el reconocimiento de objetos dado que pueden ocultarlos.

La oclusión si es un problema del reconocimiento de objetos que consta de que parte del objeto este oculto en la imagen.

### **En cuanto a las características en rectángulo, sabemos que su valor se calcula como $\Sigma$ (píxeles en área negra) - $\Sigma$ (píxeles en área blanca) ¿Pero el número de píxeles en estas dos áreas es el mismo?**

{

~ Si, para todo tipo.

~ Si ya que hay de dos tipos, dos y cuatro rectángulos, y puesto que dichos rectángulos son del mismo tamaño el número de píxeles sería igual.

= No, hay tres tipos: dos, tres y cuatro-rectángulos. Por tanto cuando tenemos tres rectángulos habría dos partes de un color y una del otro, siendo un área el doble que la otra.

}

Explicación: Como podemos observar en la diapositiva número 13 del tema, hay 3 tipos: dos, tres y cuatro-rectángulos:

Es por esto que la única respuesta correcta sería la tercera.

Nota: creo que esta pregunta fue planteada en clase, pero dudo puesto que lo único que se dice es que "Cada característica puede tener cualquier tamaño y orientación" pero no indica que los rectángulos blancos y negros tengan que ser del mismo tamaño. ¿Es así?

**Hablando sobre The Attentional Cascade, los autores proponen una heurística y una de las características es:{**

= El usuario fija tanto el nivel de falsos positivos aceptables como el ratio de detección.

~ No le afecta el cambio de luminosidad.

~ Sólo sirve para caras frontales o con poco giro.

}

**En el reconocimiento de caras, ¿Qué usamos en la extracción de características? {**

~ La imagen original, tal cual.

= Una imagen integral.

~ Ninguna de las anteriores.

}

*Explicación: En la extracción de características en el reconocimiento de caras utilizamos una imagen integral. Esta se obtiene cogiendo la imagen original y pasándola a imagen integral, y a partir de ahí se sacan las características en tiempo constante, algo que es muy importante.*

**Sobre el reconocimiento con características, ¿Que afirmación es correcta?:{**

~ Solo extraemos los SIFT de una imagen por lo que no encontramos correspondencias entre un modelo y una imagen, no hace falta.

= Primero extraemos las características SIFT de una imagen "modelo" por lo tanto el objeto ahora es representado por sus características SIFT, luego tenemos una nueva imagen "escena" donde queremos buscar ese objeto, extraemos los SIFT de la nueva imagen y encontramos las correspondencias entre modelo y la nueva imagen.

~ Ninguna de las anteriores es correcta.

}

Explicación: La segunda afirmación es la correcta. Diapositiva 5, Tema 13.

**En la siguiente imagen del cálculo de la imagen integral, ¿qué representa  $s(x,y)$ ?**

$$s(x, y) = s(x, y-1) + i(x, y)$$

$$ii(x, y) = ii(x-1, y) + s(x, y)$$

{

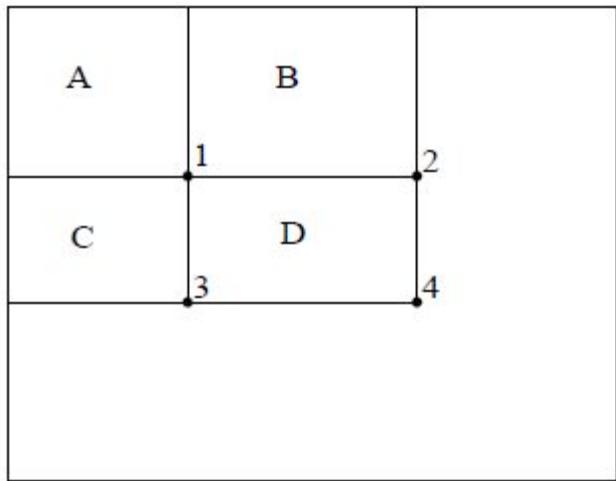
= suma acumulada de la columna

~ suma acumulada de la fila

~ ninguna de las anteriores

}

Explicación: el cálculo de la imagen integral que permite hacer el cálculo en un solo paso es el que se muestra en la imagen anterior, donde  $s(x,y)$  es la suma acumulada de la columna.



**A partir de la imagen integral representada por la figura anterior, ¿cómo se puede calcular la suma de los píxeles dentro del rectángulo B?**{

$$\sim B = II(2)$$

$$\sim B = II(4) - II(1) - II(3)$$

$$= B = II(2) - II(1)$$

}

Explicación:

Con  $II(2)$  se calcula la suma de los píxeles dentro de los rectángulos A y B. Para calcular solo la suma del rectángulo

**Dada la siguiente imagen:**

0      2      2      3

2      5      5      4

2      5      3      3

3      4      3      1

Calcular su imagen integral: {

= 0      2      4      7

  2      9      16     23

  4      16     26     36

  7      23     36     47

~ 0      2      4      7

  2      9      16     23

  4      16     23     36

  7      23     36     57

~ 0      2      4      7

  2      9      18     29

  4      18     39     71

7      29      71      143

}

Explicación: A partir de la imagen original se calcula el valor en un punto de la imagen integral sumando todos los pixeles que se sitúen arriba y a la izquierda en la imagen original más el valor del punto en la imagen original. Por ejemplo:

Siendo la imagen original (i) la dada en el enunciado, el valor del punto (2,2) en la imagen integral (ii) sería:

$$ii(2,2) = i(2,2) + (2,1) + (1,2) + (1,1) \rightarrow ii(2,2) = 5 + 2 + 2 + 0 = 9$$

### Dada la siguiente imagen de 3x3 por nivel de gris de sus pixeles

1    2    3  
0    1    1  
1    2    1

¿Cuál sería la imagen integral correspondiente?

{

= 1    3    6  
    1    4    8  
    2    7    12  
~ 1    3    6  
    1    4    10  
    2    6    16  
~ 1    3    6  
    1    4    10  
    3    7    12

}

Explicación: El valor de un pixel en la imagen integral es el sumatorio de todos los pixeles a la izquierda y arriba más el valor del propio pixel y se puede calcular en un solo paso mediante las ecuaciones siguientes.

$$s(x, y) = s(x, y-1) + i(x, y)$$

$$ii(x, y) = ii(x-1, y) + s(x, y)$$

Donde 's' representa la suma acumulada por columnas, 'i' el valor de la imagen e 'ii' el de la imagen integral en dicho punto.

Se puede comprobar así, por ejemplo, que el valor  $ii(3, 2)$  es 8 y no 10 como aparece en las dos respuestas incorrectas.

$$s(3, 2) = s(3, 1) + i(3, 2) = 3 + 1 = 4$$

$$ii(3, 2) = ii(2, 2) + s(3, 2) = 4 + 4 = 8$$

### Hablando de "The Attentional Cascade" ¿Qué característica de la heurística es FALSA? {

= Cada nivel aumenta el ratio de falsos positivos y el ratio de detección.

~ Para definir este mecanismo, hay que determinar: el número de niveles, el numero de características en cada nivel y el umbral de cada clasificador.

~ Se añaden niveles hasta alcanzar el valor tanto de falsos positivos como de ratio de detección

}

Explicación: Como vemos en la página 25 del tema 31 Vemos los términos heurísticos propuestos por los autores de dicho esquema/algoritmo. Vemos que la respuesta correcta en realidad es la siguiente:

Cada nivel reduce el ratio de falsos positivos y decrece el ratio de detección.

**Dada la siguiente matriz de una imagen, señala la imagen integral correspondiente:**

3 7 9 12 13

1 3 4 8 10

2 9 1 0 23

13 7 8 15 11

9 6 1 11 17

{

~

3 10 19 21 34

4 14 27 47 70

6 25 39 59 105

19 45 67 102 159

28 60 83 129 208

~

3 10 19 21 34

4 14 27 47 70

6 25 39 59 105

19 45 67 102 159

28 60 83 124 203

= Ninguna es correcta

}

1 0 1 2

0 1 2 3

1 2 1 4

2 3 4 1

**Dado la matriz arriba indicada de una imagen, cual de las siguientes matrices describe su imagen integral correspondiente.**

{

~

1	1	2	4
1	5	7	10
2	13	14	18
4	23	27	28

=

1	1	2	4
1	2	5	10
2	5	9	18
4	10	18	28

~

1	1	2	4
1	2	5	10
11	5	9	18
4	10	18	28

}

**Dada la siguiente imagen en escala de grises de tamaño 5x5,**

0	1	2	3	4
1	1	2	3	4
2	1	2	3	4
3	1	2	3	4
4	1	2	3	4

y las matrices:

X1

0	1	3	6	10
1	3	7	13	21
3	6	12	21	33
6	10	18	30	46
10	15	25	40	60

X2

0	1	3	6	10
1	3	7	15	21
3	6	12	21	33
6	10	18	30	46
10	15	25	40	60

X3

0	1	3	6	10
1	3	7	13	21
3	6	12	21	33
6	12	18	30	46
10	15	25	40	60

Indica cual de las tres sería su Imagen Integral: {

- = X1
- ~ X2
- ~ X3
- }

Dada la siguiente imagen en escala de grises de tamaño 5x5,

0	1	2	3	4
1	1	2	3	4
2	1	2	3	4
3	1	2	3	4
4	1	2	3	4

y las matrices:

X1

0	1	3	6	10
1	3	7	13	21
3	6	10	19	31
6	10	16	28	44
10	15	23	38	58

X2

0	1	3	6	10
1	3	7	13	21
3	6	12	21	33
6	10	18	30	46
10	15	25	40	60

X3

0	1	3	6	10
1	3	7	13	21
3	6	14	23	35
6	10	20	32	48
10	15	27	42	62

Indica cual de las tres sería su Imagen Integral: {

- ~ X1
- = X2
- ~ X3
- }

**Dada la siguiente imagen en escala de grises de tamaño 5x5,**

0	1	2	3	4
1	1	2	3	4
2	1	2	3	4
3	1	2	3	4
4	1	2	3	4

y las matrices:

X1

0	1	3	6	10
1	3	7	13	21
3	6	12	21	33
6	13	21	33	49
10	18	28	43	63

X2

0	1	3	6	10
1	3	7	10	18
3	6	12	18	30
6	10	18	27	43
10	15	25	37	57

X3

0	1	3	6	10
1	3	7	13	21
3	6	12	21	33
6	10	18	30	46
10	15	25	40	60

Indica cual de las tres sería su Imagen Integral: {

~ X1

~ X2

= X3

}

**Dada la siguiente imagen en escala de grises de tamaño 5x5,**

0	1	2	3	4
1	1	2	3	4
2	1	2	3	4
3	1	2	3	4
4	1	2	3	4

y las matrices:

X1

0	1	3	6	10
1	3	7	13	21
3	6	12	21	35
6	10	18	30	48
10	15	25	40	62

X2

0	1	3	6	10
1	3	7	13	21

3	6	12	21	33
6	10	18	30	46
10	15	25	40	60

X3

0	1	3	6	10
1	3	7	13	21
4	7	13	22	34
6	10	18	30	46
10	15	25	40	60

Indica cual de las tres sería su Imagen Integral: {

- ~ X1
- = X2
- ~ X3
- }