

# Tareas

martes, 4 de marzo de 2025

08:05 p. m.

# Juego de imitación imitado

viernes, 17 de enero de 2025 01:11 p. m.

IA@UNISON NotebookLM

[https://notebooklm.google.com/notebook/b848bc74-0545-45fc-b060-b44c472567ec?\\_gl=1\\*1wtn1q7\\*\\_ga\\*MTY5MzMwMTA4NS4xNzM3Mzg4Nzk4\\*\\_ga\\_W0LDH41ZCB\\*MTczNzM4ODc5OC4xLjAuMTczNzM4ODgwOC41MC4wLjA.&original\\_referer=https%3F%2Fnotebooklm.google%23](https://notebooklm.google.com/notebook/b848bc74-0545-45fc-b060-b44c472567ec?_gl=1*1wtn1q7*_ga*MTY5MzMwMTA4NS4xNzM3Mzg4Nzk4*_ga_W0LDH41ZCB*MTczNzM4ODc5OC4xLjAuMTczNzM4ODgwOC41MC4wLjA.&original_referer=https%3F%2Fnotebooklm.google%23)

## Resumen

El texto es un artículo de Alan Turing que explora la pregunta "¿Pueden pensar las máquinas?". Turing propone el "juego de la imitación" como una prueba para la inteligencia artificial, donde un interrogador debe distinguir entre un humano y una máquina a través de respuestas escritas. El artículo analiza diversas objeciones a la posibilidad de máquinas pensantes, incluyendo objeciones teológicas, matemáticas, y basadas en la conciencia. Finalmente, Turing argumenta a favor de la creación de "máquinas de aprendizaje", inspiradas en el desarrollo infantil y la evolución biológica, como la vía más prometedora para lograr una inteligencia artificial capaz de pasar su prueba.

El artículo explora la naturaleza de la inteligencia, la posibilidad de que las máquinas puedan pensar, y responde a diversas objeciones a esta idea. Turing no pretende definir el pensamiento, sino proporcionar un criterio operativo para determinar si una máquina puede exhibir un comportamiento indistinguible del de un ser humano.

## Temas Principales:

El Juego de Imitación (o Test de Turing):

Reformulación de la Pregunta: Turing abandona la pregunta abstracta "¿Pueden pensar las máquinas?" y la reemplaza por una más concreta: "¿Puede una máquina imitar a un ser humano en una conversación hasta el punto en que un interrogador no pueda distinguirla?".

Descripción del Juego: El juego involucra a un interrogador (C) y dos sujetos, un hombre (A) y una mujer (B), en habitaciones separadas. El interrogador debe determinar quién es el hombre y quién es la mujer mediante preguntas escritas. Turing propone reemplazar al hombre (A) con una máquina para ver si el interrogador puede distinguir entre la máquina y la mujer.

Propósito del Juego: El objetivo no es imitar perfectamente, sino persuadir al interrogador de que es un ser humano.

Cita: "En lugar de intentar tal definición, reemplazaré la pregunta por otra, que está estrechamente relacionada con ella y se expresa en palabras relativamente inequívocas."

Crítica del Nuevo Problema:

Ventajas del Test: El juego de imitación enfoca la atención en las capacidades intelectuales en lugar de las físicas, al impedir que el interrogador vea o toque a los competidores.

Ejemplos de Interacción: Turing proporciona ejemplos de preguntas y respuestas que podrían surgir en el juego, mostrando que se puede incluir casi cualquier campo del conocimiento humano.

Objeción de Desigualdad: Se plantea la preocupación de que el juego favorezca a los humanos, pero Turing señala que si una máquina puede jugar satisfactoriamente, esta objeción se vuelve irrelevante.

Cita: "El nuevo problema tiene la ventaja de trazar una línea bastante clara entre las capacidades físicas e intelectuales de un hombre."

Las Máquinas en el Juego:

Restricción a Computadoras Digitales: Turing limita la participación en el juego a computadoras digitales, argumentando que esto es una restricción práctica, no una limitación fundamental.

Justificación: Se enfoca en las computadoras digitales debido a que eran el tipo de máquinas que más interés habían despertado en ese momento.

Cita: "Esta restricción parece a primera vista ser muy drástica. Intentaré mostrar que en realidad no lo es."

Computadoras Digitales:

Modelo del Computador Humano: La idea detrás de las computadoras digitales es que imitan las operaciones que haría un "computador humano" que sigue reglas fijas.

Componentes: Describe las tres partes principales de una computadora digital: la memoria (store), la unidad de procesamiento (executive unit) y el control.

Programación: Explica que programar una máquina implica convertir las reglas del computador humano en una tabla de instrucciones que la máquina pueda entender.

Computadoras con Elemento Aleatorio: Discute la posibilidad de computadoras digitales con un elemento aleatorio que podría dar la impresión de "libre albedrío".

Cita: "La idea detrás de las computadoras digitales se puede explicar diciendo que estas máquinas están destinadas a llevar a cabo cualquier operación que pudiera realizar un computador humano."

Universalidad de las Computadoras Digitales:

Máquinas de Estado Discreto: Explica que las computadoras digitales pertenecen a la categoría de máquinas de estado discreto, las cuales se mueven de un estado a otro de forma abrupta y definida.

Capacidad de Imitación: Las computadoras digitales tienen la capacidad de imitar cualquier máquina de estado discreto, siempre que tengan suficiente velocidad y capacidad de almacenamiento.

Máquinas Universales: Describe a las computadoras digitales como "máquinas universales" debido a su capacidad para realizar cualquier cálculo.

Cita: "Esta propiedad especial de las computadoras digitales, que pueden imitar cualquier máquina de estado discreto, se describe diciendo que son máquinas universales."

Visiones Contradictorias Sobre la Pregunta Principal:

Predicciones de Turing: Turing predice que en 50 años las computadoras podrán jugar al juego de la imitación tan bien que un interrogador promedio tendrá solo un 70% de probabilidad de acertar tras 5 minutos de preguntas.

Objeciones: Se presentan y refutan varias objeciones a la idea de máquinas pensantes:

Objeción Teológica: La idea de que el pensamiento es exclusivo del alma inmortal humana. Turing responde sugiriendo que esto limita el poder de Dios.

Objeción del "Esconder la Cabeza en la Arena": El temor a las consecuencias de que las máquinas piensen. Turing señala que esta objeción no tiene fundamento racional.

Objeción Matemática: La limitación del poder de las máquinas debido al teorema de Gödel y otros resultados lógicos. Turing argumenta que la inteligencia humana también puede estar sujeta a limitaciones.

Argumento de la Conciencia: Que las máquinas no pueden sentir emociones o comprender el significado de sus acciones. Turing la califica como solipsista e innecesaria para el propósito del juego.

Argumentos de Diversas Discapacidades: La afirmación de que las máquinas no pueden ser creativas, cometer errores, etc. Turing responde que muchas de estas limitaciones son el resultado de una falta de capacidad de almacenamiento.

Objeción de Lady Lovelace: Que las máquinas solo pueden hacer lo que se les ordena. Turing argumenta que las máquinas pueden sorprendernos y que los humanos pueden haber originado sus ideas gracias a la educación.

Argumento de la Continuidad en el Sistema Nervioso: Que el sistema nervioso es continuo y las máquinas son discretas. Turing argumenta que la prueba del juego de la imitación no requiere imitar la estructura física del sistema nervioso.

Argumento de la Informalidad del Comportamiento: Que el comportamiento humano no puede ser descrito por reglas. Turing responde que aunque puede que no haya reglas de conducta definidas, existen leyes de comportamiento que rigen a las personas.

Argumento de la Percepción Extrasensorial (ESP): Turing reconoce la fuerza de este argumento, pero señala que se puede mejorar el juego para evitarlo.

Cita: "La pregunta original, '¿Pueden pensar las máquinas?', creo que es demasiado sin sentido para merecer discusión."

Máquinas Aprendices:

Aprendizaje en lugar de Programación Directa: Turing propone crear máquinas que puedan aprender en lugar de ser programadas directamente para comportarse como adultos.

Analogía con el Niño: Sugiere simular el cerebro de un niño y someterlo a un proceso de educación.

Analogía con la Evolución: Compara el proceso de experimentación y mejora de las máquinas aprendices con la evolución biológica.

Importancia de la Comunicación: Enfatiza la necesidad de canales de comunicación "no emocionales" para la enseñanza de la máquina.

Inclusión del Elemento Aleatorio: Turing sugiere que incluir elementos aleatorios puede ser beneficioso para el aprendizaje, tal como lo hace la evolución biológica.

Cita: "En el proceso de intentar imitar una mente humana adulta, estamos obligados a pensar mucho en el proceso que la ha llevado al estado en el que se encuentra."

Ideas Clave y Conclusiones:

El Test de Turing como Criterio Operativo: Turing propuso el juego de la imitación como una forma práctica de explorar si una máquina puede comportarse de manera inteligente, en lugar de definir la inteligencia.

Énfasis en el Comportamiento, No en la Esencia: El foco está en la capacidad de una máquina para exhibir un comportamiento inteligente, no en si posee o no un "alma" o conciencia.

Universalidad de las Computadoras Digitales: Las computadoras digitales son máquinas poderosas capaces de simular cualquier máquina de estado discreto.

Posibilidad del Aprendizaje en las Máquinas: La idea de que las máquinas pueden aprender y evolucionar mediante la interacción con un "profesor", lo que elimina la necesidad de programación directa.

Naturaleza Conjetural de las Ideas: Turing reconoce que muchas de sus afirmaciones son conjeturas, pero destaca la importancia de explorar estas ideas para impulsar la investigación.

Impacto y Relevancia:

El artículo de Turing no solo introdujo el concepto del Test de Turing, sino que también sentó las bases para la investigación en inteligencia artificial y sigue siendo relevante para los debates actuales sobre la naturaleza de la inteligencia y las capacidades de las máquinas. Su enfoque en el aprendizaje y la evolución sigue siendo un tema central en el campo de la IA.

## Preguntas Frecuentes

¿Cuál es el problema original que Turing busca reformular, y por qué lo hace?

El problema original que Turing busca reformular es la pregunta: "¿Pueden las máquinas pensar?" Turing considera que esta pregunta es demasiado vaga y está cargada de definiciones ambiguas de términos como "máquina" y "pensar". En lugar de tratar de definirlos, prefiere reemplazar la pregunta por una más precisa que pueda abordarse experimentalmente, evitando la necesidad de un estudio estadístico sobre el uso común de las palabras.

¿En qué consiste el "Juego de Imitación" y cuál es su propósito en el contexto del pensamiento de las máquinas?

El "Juego de Imitación" involucra a tres participantes: un hombre (A), una mujer (B) y un interrogador (C). El interrogador está separado de los otros dos y su objetivo es determinar cuál es el hombre y cuál la mujer mediante preguntas escritas. Luego, Turing propone que una máquina tome el lugar del hombre (A). La pregunta clave se convierte en: "¿Puede la máquina engañar al interrogador con tanta frecuencia como lo haría un hombre imitando a una mujer?" El propósito de este juego es ofrecer un criterio tangible y medible para determinar si una máquina puede exhibir un comportamiento que podría considerarse como "pensamiento".

¿Por qué Turing limita el juego a las computadoras digitales, y cómo define él estas máquinas?

Turing limita el juego a las computadoras digitales porque son el tipo de máquina que ha generado el interés en las "máquinas pensantes". Descarta otras formas de máquinas por consideraciones prácticas y para enfocar el debate. Describe una computadora digital como una máquina que puede llevar a cabo cualquier operación que un computador humano podría realizar siguiendo un conjunto de reglas fijas, que incluyen una unidad de almacenamiento, una unidad ejecutiva y una unidad de control. Estas unidades corresponden a la memoria, el procesador y el controlador del flujo de las operaciones en un computador humano.

¿Qué son las "máquinas de estados discretos" y qué importancia tienen en la argumentación de Turing sobre la posibilidad de la inteligencia en máquinas?

Las "máquinas de estados discretos" son aquellas que se mueven entre estados definidos y distintos, como un interruptor de luz (encendido o apagado). Turing clasifica las computadoras digitales dentro de esta categoría, aunque con una cantidad de estados extremadamente grande. Su importancia radica en que se puede predecir su comportamiento conociendo su estado inicial y la secuencia de señales de entrada. Además, las computadoras digitales pueden imitar cualquier máquina de estados discretos si tienen suficiente capacidad de almacenamiento y velocidad, lo que las convierte en "máquinas universales". Esta característica es crucial para la argumentación sobre si una máquina puede imitar al cerebro humano.

¿Cuáles son las principales objeciones a la posibilidad de máquinas pensantes que Turing discute en su artículo,

y cómo las refuta?

Turing aborda varias objeciones, incluyendo:

La objeción teológica: Que el pensamiento es una función del alma humana dada por Dios. Turing la rechaza argumentando que limitaría la omnipotencia divina.

La objeción "cabeza en la arena": El temor de que las máquinas pensantes sean una amenaza. Turing no le da peso y sugiere que esta emoción debería ser reemplazada por la calma.

La objeción matemática: Que los resultados de la lógica matemática muestran limitaciones en las máquinas.

Turing responde que estas limitaciones son también posibles en el intelecto humano.

El argumento de la conciencia: Que las máquinas no pueden tener las mismas emociones y conciencia que los humanos. Turing destaca el problema del solipsismo que surge de esta objeción y ofrece el juego de imitación como una manera de superar este problema.

Las objeciones basadas en diversas discapacidades: Que las máquinas nunca serán capaces de hacer ciertas cosas como ser amables o creativas. Turing argumenta que estas son limitaciones del estado actual de las máquinas y sus capacidades de almacenamiento.

La objeción de Lady Lovelace: Que las máquinas solo pueden hacer lo que les decimos. Turing responde que con el aprendizaje de las máquinas esto puede ser superado.

El argumento de la continuidad del sistema nervioso: Que un sistema de estados discretos no puede replicar un sistema continuo como el sistema nervioso. Turing señala que en el juego de imitación esta diferencia no es relevante.

El argumento de la informalidad del comportamiento: La dificultad de programar reglas de comportamiento generales. Turing concuerda con este problema pero argumenta que la ausencia de reglas de conducta no implica la ausencia de leyes de comportamiento que puedan regir el comportamiento de las máquinas.

El argumento de la percepción extrasensorial (ESP): Que la ESP podría ser un obstáculo insalvable en el juego.

Turing está de acuerdo con la solidez de esta objeción pero no le otorga mucha importancia, ya que este problema sería el mismo con una persona.

¿Qué implicaciones tiene el concepto de "máquinas universales" en la visión de Turing sobre las máquinas pensantes?

El concepto de "máquina universal", que puede imitar el comportamiento de cualquier otra máquina de estados discretos, significa que no se necesitan diferentes máquinas para realizar diferentes tipos de tareas de computación. Una sola computadora digital, con la programación adecuada, puede hacer cualquier cosa que cualquier otra computadora haga. Esto implica que la diferencia entre los tipos de máquinas se reduce a una cuestión de programación y capacidad de almacenamiento, y no a limitaciones inherentes en la tecnología de las computadoras. Así pues, la misma máquina puede ser usada para "pensar", siendo una forma de computación.

¿Cómo se plantea Turing el desarrollo de "máquinas de aprendizaje" y qué similitudes encuentra entre este proceso y la evolución biológica?

Turing propone que, en lugar de programar directamente una máquina que imite a un adulto, se debe desarrollar una "máquina infantil" que luego se eduque. Este enfoque imita la forma en que un ser humano aprende y se desarrolla. Turing ve similitudes entre el proceso de aprendizaje de una máquina y la evolución biológica, comparando la estructura de la máquina infantil con el material hereditario, las modificaciones de la máquina con las mutaciones, y la selección por el experimentador con la selección natural.

¿Cuál es la predicción de Turing sobre el futuro de las máquinas pensantes y qué pasos sugiere para alcanzarla?

Turing predice que alrededor del año 2000, será posible programar computadoras para que jueguen el "Juego de Imitación" de manera tan convincente que un interrogador promedio no podrá distinguir la máquina de un ser humano en más del 70% de las veces después de cinco minutos de conversación. Para lograrlo, sugiere enfocarse en mejorar la programación, desarrollar "máquinas infantiles" y perfeccionar los procesos de enseñanza y aprendizaje, incluyendo el uso de recompensas y castigos, así como la comunicación a través de lenguajes simbólicos. Turing también indica que se necesita espacio de almacenamiento, pero que la velocidad no será un problema.

# Checklist para el Desarrollo Ético de Sistemas de Inteligencia Artificial

martes, 4 de febrero de 2025 05:13 p. m.

1. Definición y Propósito
  - Establecer claramente el objetivo y las aplicaciones del algoritmo, asegurando su alineación con fines legítimos y seguros, y considerando la población objetivo.
2. Colaboración con Expertos
  - Consultar a especialistas en el área para validar el diseño y uso de la IA, y documentar sus recomendaciones.
3. Seguridad y Privacidad
  - Seguridad desde el diseño (SdB):** Identificar y mitigar riesgos potenciales, garantizando ciberseguridad y resistencia frente a fallos o ataques.
  - Privacidad de los datos:** Cumplir con la normativa vigente (por ejemplo, GDPR y leyes locales) mediante evaluaciones de impacto, minimización y protección de datos personales, almacenándolos solo el tiempo necesario.
4. Trazabilidad y Documentación
  - Mantener registros completos de las acciones, decisiones y datos utilizados en el desarrollo, de modo que se pueda auditar y replicar el proceso.
5. Control, Transparencia y Evaluación
  - Garantizar que los procesos de toma de decisiones sean comprensibles y supervisados por humanos, permitiendo la intervención en caso de errores o riesgos que afecten la integridad física o psicológica de las personas.
  - Realizar evaluaciones periódicas para asegurar que la IA opere de forma legítima y sin impactos discriminatorios.
6. Confiabilidad, Robustez e Inclusión
  - Probar el sistema en entornos que simulen situaciones reales, asegurando resultados fiables y la capacidad de manejar condiciones adversas.
  - Promover la accesibilidad y usabilidad, considerando posibles prejuicios o limitaciones que puedan afectar a sectores vulnerables.
7. Responsabilidad y Mejora Continua
  - Definir responsabilidades claras en el desarrollo y supervisión del sistema, con un organismo interno que garantice la ética y seguridad del uso de la IA.
  - Establecer mecanismos de seguimiento y retroalimentación para adaptar el sistema a nuevas circunstancias y mejorar continuamente su desempeño.

**Nota personal:**  
El desarrollo de la inteligencia artificial debe garantizar siempre la seguridad y la privacidad de las personas. Ningún sistema de IA debe diseñarse para causar daño de ninguna forma, y su implementación debe cumplir con principios de ciberseguridad para prevenir fallos o ataques que puedan poner en riesgo la integridad de los usuarios. Además, la IA debe ser confiable y explicable. Su proceso de toma de decisiones debe ser comprensible y auditable para evitar errores y sesgos que puedan generar daños a la integridad de las personas.

Sin embargo, aunque la regulación sea necesaria, esta no debe entorpecer la innovación. Es posible equilibrar la seguridad con el avance tecnológico mediante principios de responsabilidad y transparencia, asegurando que la IA se desarrolle de forma ética y para beneficio de la sociedad.

Primer borrador:

1. Definición y Propósito
  - Definir el propósito y posibles usos del algoritmo, incluyendo la población objetivo y el servicio ofrecido.
  - Asegurar que la función objetivo del sistema esté totalmente alineada con su propósito y considerar problemas de seguridad que pueda tener.
2. Colaboración con Expertos
  - Consultar con especialistas en las áreas de aplicación del algoritmo.
  - Documentar las recomendaciones y advertencias de los expertos.
3. Seguridad
  - Implementar el principio de seguridad desde el diseño (SdB), evaluando riesgos potenciales e implementando prevenciones desde el inicio.
  - Evaluar la resistencia del sistema ante errores, fallos e incoherencias, así como su ciberseguridad.
  - Establecer medidas para resistir errores y fallos, especialmente en la interacción con personas u otros sistemas.
4. Trazabilidad
  - Mantener un registro de acciones y datos utilizados durante todo el proceso de desarrollo.
  - Asegurar la comprensión y replicación del desarrollo del sistema.
5. Evaluación y Regulación
  - Evaluar, regular y asegurar un control humano suficiente sobre los algoritmos, especialmente a medida que ganan poder y autonomía.
  - Garantizar que las personas afectadas por el sistema estén protegidas si las decisiones automatizadas les perjudican.
6. Principios Éticos
  - Asegurar el trato a las personas con el debido respeto.
  - Promover la dignidad humana y evitar todo tipo de daño, tanto físico como psicológico.
  - Asegurar que el uso de la IA se alinee con los principios de integridad académica: honestidad, respeto, confianza, responsabilidad, justicia y valentía.
  - Utilizar la IA teniendo en cuenta el bien común, respetando la dignidad de las personas y promoviendo el desarrollo sostenible.
  - Estimar el daño en caso de fallo del sistema, considerando gravedad, escala y probabilidad.
7. Transparencia y Explicabilidad
  - Garantizar que el sistema sea comprensible por un ser humano, desde la captación de datos hasta la toma de decisiones.
  - Asegurar que la persona tenga control informado sobre la actividad del sistema de IA.
  - Implementar una obligación de información a los interesados y registro de la documentación relativa al sistema de IA.
  - Permitir la supervisión humana y la capacidad de impugnar decisiones automatizadas.
8. No Discriminación y Justicia
  - Asegurar la justicia en los resultados de la IA, sin impactos discriminatorios.
  - Asegurar que los datos que alimentan las soluciones de IA sean exactos, no sesgados y lícitos, justos y transparentes.
  - Implementar mecanismos de supervisión, gobernanza y gestión de datos.
9. Control y Vigilancia Humana
  - Garantizar la presencia humana en algún momento del proceso de toma de decisiones automatizadas.
  - Asegurar que las personas que interactúan con el sistema sean conscientes de que se trata de una IA.
  - Establecer controles para medir la contribución de la IA al bienestar humano, asegurando que sus beneficios superen sus costos.
  - Identificar al personal que revisa las decisiones y verificar los resultados emitidos por el sistema.
  - Permitir la intervención humana o la interrupción del funcionamiento del sistema.
  - Invalidar o revertir la información de salida generada por el sistema.
10. Privacidad y Protección de Datos
  - Asegurar la protección de datos personales y cumplir con la normativa vigente.
  - Realizar una evaluación del impacto de las operaciones de tratamiento en la protección de datos personales, si estas suponen un alto riesgo.
  - Diseñar sistemas de IA para mantener la integridad de los datos personales, usándolos solo el tiempo necesario para los fines definidos.
11. Confiabilidad y Robustez
  - Asegurar que los sistemas de IA funcionen de forma fiable, segura y coherente en circunstancias normales y adversas.
  - Garantizar que los resultados de los sistemas sean confiables y reproducibles en cualquier situación.
  - Probar el sistema en entornos que simulen situaciones del mundo real.
  - Gestionar ataques y defensas contra modelos inteligentes.
12. Inclusión
  - Promover la accesibilidad y usabilidad para personas con discapacidad.
  - Considerar problemas cognitivos, prejuicios sociales o déficits en el aprendizaje.
13. Responsabilidad
  - Establecer responsabilidades claras para todas las personas involucradas en el desarrollo de la IA.
  - Establecer un organismo de revisión interno para supervisar y orientar el desarrollo y la implementación de sistemas de IA.
  - Capacitar al personal en el uso y mantenimiento responsable y ético de la solución.
14. Seguimiento y Mejora Continua
  - Supervisar y evaluar el rendimiento del sistema de forma continua.
  - Recopilar feedback de los usuarios para mejorar el sistema y asegurar el cumplimiento de los principios éticos.
  - Evaluar el impacto de las nuevas aplicaciones del sistema en las personas.
15. Documentación y Comunicación
  - Declarar y referenciar el uso de IA en todo momento.
  - Asegurar que la información facilitada sea accesible y comprensible.
  - Documentar todos los pasos del desarrollo, incluyendo datos, decisiones y evaluaciones.

Este checklist está diseñado para ser un punto de partida. Es importante adaptar los principios a las particularidades de cada proyecto y sector, y siempre tener en cuenta que el desarrollo de la IA es un campo en constante evolución.

# Aprendizaje-arboles de decisión

jueves, 6 de febrero de 2025 11:13 a. m.

1. arboles\_cualitativos.py y que luego se utilizan en prueba\_arbol\_Q.py. Para eso, vamos a buscar algunas bases de datos sencillas que permitan ilustrar cómo funciona el algoritmo de aprendizaje, así como su uso en predicción. Estos conjuntos de datos se obtuvieron de UC Irvine Machine Learning Repository
2. arboles\_numericos.py y se puede ver como utilizarlo en prueba\_arbol\_N.py, también con un conjunto de datos del mismo UC Irvine Machine Learning Repository.

Para esta tarea, desarrollar y probar un algoritmo de *bosques aleatorios* basado en los algoritmos ya presentados, hay que hacer lo siguiente:

1. En el archivo arboles\_numericos.py se encuentra la siguiente función:

```
def entrena_arbol(datos, target, clase_default,
                  max_profundidad=None, acc_nodo=1.0, min_ejemplos=0,
                  variables_seleccionadas=None):
```

En el cual existe el parámetro variables\_seleccionadas el cual no se usa (por default usa el valor None). Modifica esta función (no tienes que modificar en otras subfunciones o clases) de manera que si el valor de este parámetro es un número entero, entonces se seleccione solo ese número de atributos, seleccionados aleatoriamente, para seleccionar el par (atributo, valor) en ese nodo (solo en ese nodo). Recuerda que a los hijos se les envía todos los atributos en forma recursiva.

2. En el archivo bosque\_aleatorio.py desarrolla una función para entrenar busques aleatorios, basándose en el uso de la función entrena\_arbol del modulo arboles\_numericos.py. El algoritmo deberá:
  1. Separar los datos en subconjuntos con selección aleatoria con repetición (para M subconjuntos).
  2. Por cada subconjunto, entrenar un árbol con un número limitado de variables en cada nodo. Un bosque se puede representar como una lista de nodos raíz (árboles).
  3. Una función para poder hacer predicciones a partir del bosque (lista de nodos raíz). Puedes usar varias funciones, que separen un problema grande en varios más pequeños y tratables.
3. En el archivo prueba\_BA.py usa un conjunto de datos **diferente** a los que ya usamos para árboles cualitativos y numérico. Muestra que pasa conforme se aumenta el número de árboles en tu bosque, la máxima profundidad de los árboles, o la cantidad de variables a considerar en cada nodo.

Este proyecto solo se puede evaluar a partir de prueba\_BA.py, por lo que es necesario desarrollar los 3 incisos y se evaluará en relación a la funcionalidad del algoritmo, y del estudio realizado en prueba\_BA.py.

# Sesgo cognitivo en aprendizaje automático (AA)

jueves, 30 de enero de 2025 01:22 p. m.

## Esencia:

- Existe un patrón, no es posible establecerlo de forma analítica.
- Tenemos datos, o los podemos generar

## Consideraciones:

- AA está permeando en el mundo laboral.
  - No sustituye al conocimiento experto. Riesgo de reducción de trabajos no especializados con el uso de AA.
- El AA depende de la evidencia, las cuales pueden tener sesgos (racismo, sexismo, etc.)
- Riesgo latente en aplicaciones no éticas.

# Puzzle un poco diferente

martes, 21 de enero de 2025 05:02 p. m.

El *rompecabezas deslizante* es una versión diferente del 15 puzzle, en la cual cada línea y cada columna se deslizan, como si se encontrara en una esfera (por supuesto que este tipo de rompecabezas no se puede hacer en madera, pero en la computadora es facilísimo). Un esquema del entorno es el siguiente:

	↑↑	↑↑	↑↑	↑↑	
←←	1	2	3	4	⇒⇒
←←	5	6	7	8	⇒⇒
←←	9	10	11	12	⇒⇒
←←	13	14	15	16	⇒⇒
	↓↓	↓↓	↓↓	↓↓	

Las acciones que el agente puede realizar sobre el ambiente son: a) Girar por la derecha el renglón  $ii$  ( $i \in 1,2,3,4; i \in 1,2,3,4$ ); b) Girar por la izquierda el renglón  $ii$  ( $i \in 1,2,3,4; i \in 1,2,3,4$ ); c) Girar por arriba la columna  $jj$  ( $j \in 1,2,3,4; j \in 1,2,3,4$ ); y d) Girar por abajo la columna  $jj$  ( $j \in 1,2,3,4; j \in 1,2,3,4$ ). Se asume que el ambiente es completamente observable. El objetivo del agente es que, después de aplicar un cierto número de movimientos aleatorios y no observados. El agente pueda realizar las acciones necesarias para regresar el sistema al estado mostrado en el esquema anterior, utilizando la menor cantidad de movimientos posibles.

## Conteste las siguientes preguntas (10 puntos por pregunta):

1. Establezca una manera de representar el estado del problema.  
Arreglo de 16 componentes (uno para cada espacio en la cuadrícula) donde puede tomar valores desde 1 a 16.
2. Establezca cuáles serían las acciones legales en un estado dado.  
Tupla (2 componentes): (izq/der/arriba/abajo, 1/2/3/4)
3. Establezca el estado sucesor a un estado dado, si se selecciona una acción.  
Dirección, ind = acción

```
Si direccion = izq
    i = 4*(ind-1)
    aux = Xi
    Para j desde 0 a 3 (excluyente)
        X[i+j] = X[i+1]
    X[i+3] = aux
```

```
Si direccion = der
    i = 4*(ind-1) + 3
    aux = Xi
    Para j desde 0 a 3 (excluyente)
        X[i-j] = X[i-1]
    X[i-3] = aux
```

```
Si direccion = arriba
    i = (ind-1)
    aux = Xi
    Para j desde 0 a 3 (excluyente)
```



```

X[i+4j] = X[i+4]
X[i+12] = aux

```

```

Si direccion = abajo
  i = (ind-1) + 12
  aux = Xi
  Para j desde 0 a 3 (excluyente)
    X[i-4j] = X[i-4]
  X[i-12] = aux

```

4. Establezca el costo local dependiendo del estado y la acción.  
Costo de 1 para cualquier acción.
5. ¿Cuál es la cardinalidad del espacio de estado?  
16!
6. ¿La distancia de Manhattan, o el número de piezas mal colocadas podrían ser heurísticas admisibles?

No, ya que si usamos por ejemplo el siguiente estado:

1	2	3	4
6	7	8	5
9	10	11	12
13	14	15	16

Vemos que el puzzle está a un movimiento de ser completado correctamente (fila 2 a la derecha), pero en la heurística de la distancia de Manhattan ( $1 + 1 + 1 + 3$ ) o el número de piezas mal colocadas (4) es mayor al costo de hacer un movimiento.

7. Desarrolle 2 heurísticas ( $h_1(n)$  y  $h_2(n)$ ) para resolver el problema por el método de búsqueda  $A^*$ .
8. Demuestre (o haga un esbozo de demostración) que las heurísticas son admisibles.

Una posible heurística sería la cantidad de veces que al recorrer el arreglo se encuentra un número menor cuando pasas de  $X_i$  a  $X_{i+1}$ . En esta heurística, el ejemplo anterior sería de 1, por ejemplo (cuando pasa de 8 a 5).

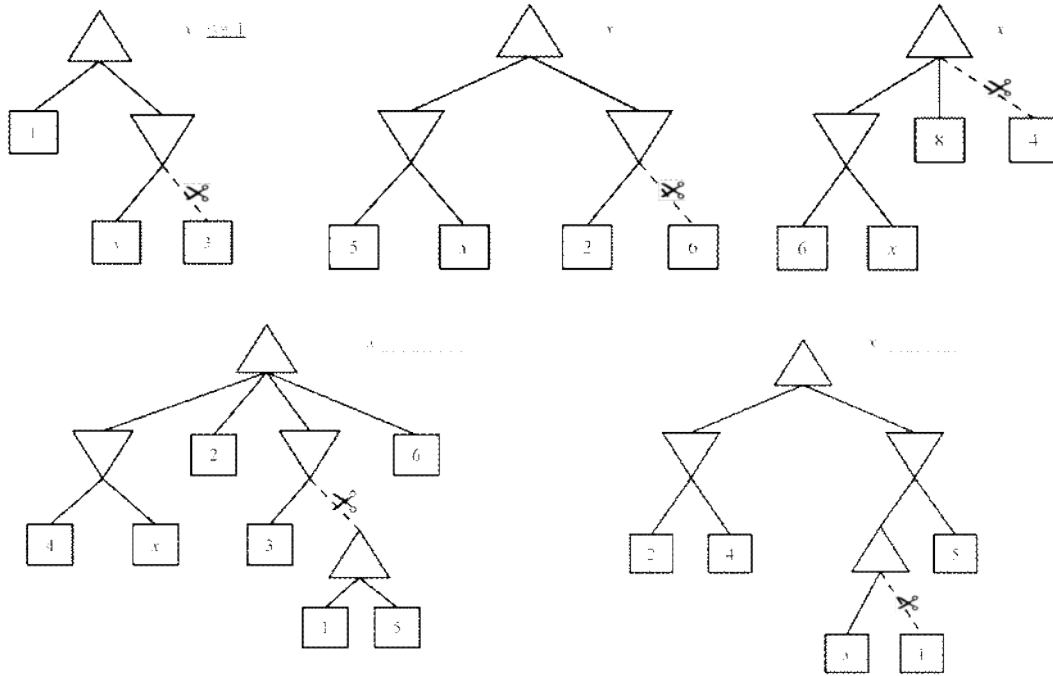
9. Determine si una heurística es dominante respecto a la otra. Demuestre que lo son (o que no lo son, en su caso).
10. ¿La búsqueda en grafos ofrecería ventajas respecto a la búsqueda en árboles en este problema? Justifique su respuesta.  
Sí. La búsqueda en árboles no guarda los estados ya evaluados, por lo que se exploraría el mismo estado varias veces (sobre todo dada la característica de ser un problema cíclico).

# Búsquedas con adversarios

lunes, 10 de marzo de 2025 09:29 a. m.

## Podando árboles

En las siguientes imágenes se busca poder realizar la poda  $\alpha$ - $\beta$  en donde se marca en cada imagen. Las jugadas están ordenadas de izquierda a derecha como es usual. Para esto es necesario dar un valor a la variable  $x$  de cada figura, tal como se muestra de ejemplo en el primer problema. En caso de no poder asignar un valor donde la poda se efectúe, favor de explicitarlo claramente.

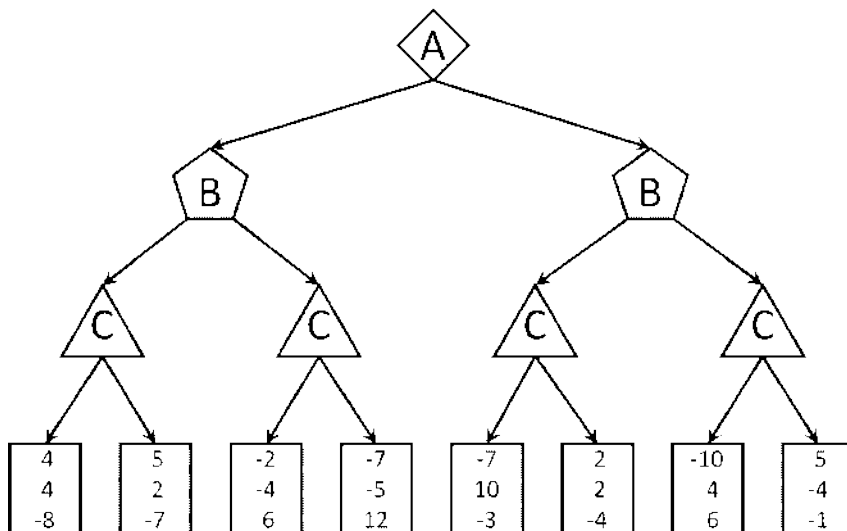


$x \geq 2$   
 $x \geq$  Hay forma?

$x \leq 4$   
 $x \leq$  Hay forma?

## Te para tres

Supongamos que tenemos un juego suma cero, pero con tres jugadores (jugadores A, B, C), los cuales solamente pueden realizar dos acciones (izquierda y derecha). Si simulamos una ronda de todos los jugadores, y los valores finales, los cuales están ordenados para el jugador A, B y C respectivamente, tenemos el árbol de juego siguiente:



1. ¿Cual es la utilidad para cada acción que pueda realizar el jugador A, si la propagamos con minimax? Escribe los valores de utilidad para cada jugada en el árbol de juego.
2. ¿Es posible hacer una poda similar a la  $\alpha$ - $\beta$  de éste árbol de juego? De ser así explica la manera y marca las ramas podadas en una poda óptima. Se no ser posible, justifica tu respuesta.

