



## Ayudantía - Search - Introducción a ML

Matías Greco, Alain Raymond  
IIC2613 - Inteligencia Artificial  
Segundo Semestre, 2019

1. Defina dos heurísticas para el problema 8-puzzle y categorícelas de acuerdo a cual es mas informada.

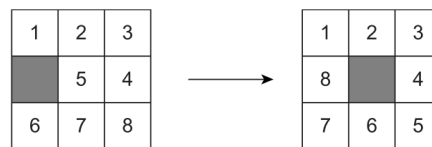


Figura 1: Ejemplo del 8-puzzle

2. Otro problema de puzzle, en el cual las piezas pueden saltar sobre otras de otro color, es el de ranas y sapos”. El objetivo es intercambiar el conjunto de cuadros blancos y negros en la mínima cantidad de movimientos. Construya como es el espacio de estados para este problema y encuentre una heurística admisible para este puzzle.

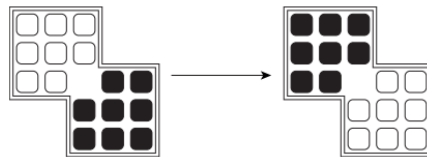


Figura 2: Ejemplo del puzzle de ranas y sapos

3. El problema del panqueque consiste en ordenar una pila de panqueques utilizando solo una espátula, la cual puede ser insertada en cualquier punto de la pila para voltear todos los panqueques encima de ella. Construya el espacio de estados y encuentre una función heurística para este problema.

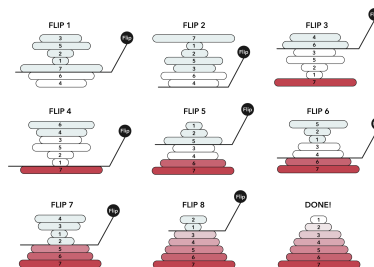


Figura 3: Ejemplo del problema del panqueque

4. En el problema anterior, ha sido demostrado teóricamente que el número máximo para ordenar  $n$  panqueques está dado por  $18 * n/11$  ¿Para que es útil conocer ese número máximo?
5. El problema del vendedor viajero (TSP) consiste en encontrar el tour de mínimo costo que recorre todas las ciudades solo una vez y vuelve a la ciudad de origen. Defina dos heurísticas para este problema, y compare su eficiencia respecto a cual expandirá menos nodos y cual hará una búsqueda mas eficiente.

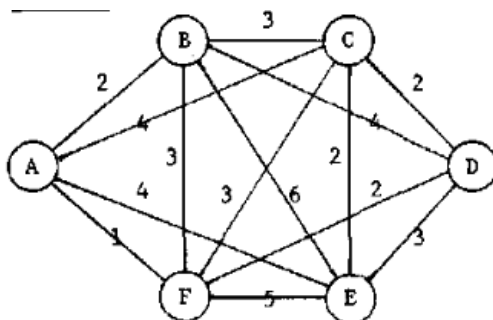


Figura 4: Espacio de búsqueda del TSP

6. ¿Por qué es necesario un conjunto de validación? ¿No sería mejor usar los datos de validación para entrenamiento?
7. ¿Hay algún contexto en que caer en *overfitting* sea bueno?
8. Verdadero o Falso: En general, la estimación del error de un clasificador en el set de entrenamiento entrega una estimación optimista del error real de este clasificador. ¿Por qué?
9. De las siguientes alternativas, ¿cuáles se deben considerar al momento de seleccionar un algoritmo de aprendizaje de máquina?:
  - a) Tipo de problema, por ejemplo, supervisado o no supervisado.
  - b) Dimensionalidad de los datos.
  - c) Número de registros con que se cuenta.
  - d) Tipo de representación necesaria para validar o interpretar la solución.
10. Tienes una empresa de seguridad y te han contado de este maravilloso software chino para reconocer caras que utiliza *Machine Learning*. Tasa de reconocimiento del 99.9%, certificado por varias autoridades chinas de distintas ciudades de ese país. Con entusiasmo, ordenas el producto y antes de venderlo, sensatamente lo pruebas. ¿Qué resultados crees que arrojarán las pruebas? ¿Por qué?
11. Si pudieses elegir qué preferirías: ¿un espacio de solución convexo con pocos datos, o un espacio de solución cóncavo con muchísimos datos?

## 1. Solution

1. Cantidad de cuadros que no están en la posición correcta. Distancia de manhattan y distancia de manhattan con conflictos lineales (dos cuadros juntos que están intercambiados, 6 y 7 en el ejemplo) .
2. Una heurística puede ser definida como seleccionar el mayor entre la distancia de manhattan de los cuadros blancos al mas cercano de los cuadros negros, o de los cuadros negros al mas cercano de los blancos.

3. Numero de panqueques no correlativos.
4. Podar (al momento de generar un nodo, si tiene un  $f$  mayor a la cota dada no agregarlo a *OPEN*) y limitar profundidad
5. MST (árbol de expansión mínima entre todas las ciudades no visitadas) e in-out heuristic (suma de las dos aristas mas cortas de cada ciudad no agregada en el subtour de modo que estas aristas no estén conectadas con nodos interiores del subtour. A esta suma se le agrega la arista mas corta del nodo inicial y final del subtour y el total se divide por dos) (Ira Pohl, The avoidance of (relative) catastrophe, heuristic competence, genuine dynamic weighting and computational issues in heuristic problem solving, 1973)