



Universidad Simón Bolívar
Departamento de Computación y Tecnología de la Información
CI5437 Inteligencia Artificial I
05 de mayo de 2014

15 - PUZZLE

Prof. Blai Bonet

Elaborado por:
Gamar Azuaje 10-10051
Wilmer Bandres 10-10055
Juan A. Escalante 10-10227

En el campo de la computación se encuentran numerosos tipos de problemas: unos más difíciles de resolver que otros. Esta diferencia de dificultad viene dada por el espacio de búsqueda de cada uno, que determina, según su amplitud, cómo se va a computar la solución, de manera que a mayor amplitud, mayor dificultad. Un ejemplo de un problema con amplio espacio de búsqueda es el 15-puzzle, un juego que trata sobre ordenar un rompecabezas de 15 piezas (que cuentan con una configuración inicial) en la menor cantidad de movimientos posible. Este problema califica como NP-Hard.

Dado el 15-puzzle como ejemplo, podemos decir que no existe un algoritmo razonablemente eficiente para resolverlo, por lo que el ser humano tiene que recurrir a algoritmos que, si bien no garantizan la eficiencia, encuentran la solución deseada. En este proyecto, los algoritmos a utilizar son A* e IDA*. Es importante aclarar sobre su uso que, aunque el tiempo de ejecución puede llegar a ser alto, con una heurística admisible bien seleccionada, se puede lograr que estos tiempos se reduzcan considerablemente.

Existen diversas heurísticas que se pueden elegir para este problema en particular, algunas de ellas son:

- La heurística de la Distancia Manhattan: una heurística admisible que basa su observación en la mínima distancia que tienen las piezas a la posición del goal, asumiendo que éstas pueden moverse sin colisionar con las demás fichas. Al final, la heurística total de una configuración del puzzle, es la suma de estas distancias.
- La heurística de la Distancia Manhattan con Linear Collision: es similar a la anterior, con la particularidad de que suma 2 unidades al total por cada dos piezas que sabemos que deben intercambiar sus posiciones para llegar a su posición final. Esto es, cuando dos piezas ya están ubicadas en su fila (o columna) objetivo, pero están al revés, entonces sabemos que al menos necesitaremos dos pasos más para poder llegar al objetivo, aunado a la Distancia Manhattan de esas piezas.

- Por último, la heurística Pattern Database (PDB): se concentra en escoger varios subconjuntos de las piezas del puzzle, tomando como piezas ignoradas a las demás y consiguiendo la mínima cantidad de pasos para colocar esas piezas en sus posiciones finales. Cuando estos subconjuntos son disjuntos, la heurística toma el nombre de Disjoint Pattern Database y la heurística como tal termina siendo la suma de las mínimas cantidades de pasos para colocar las piezas de cada patrón, ignorando a las demás. El puzzle puede ser dividido cualquier forma, formando subconjuntos de piezas de 5-5-5, 6-6-3, entre otros.

Análisis de implementación de PDB

15-puzzle 5 fichas	
Permutaciones a visitar	$16*15*14*13*12*11 = 5765760$
Número de entradas	$16*15*14*13*12 = 524160$

15-puzzle 6 fichas	
Permutaciones a visitar	$16*15*14*13*12*11*10 = 576557600$
Número de entradas	$16*15*14*13*12*11 = 5765760$

15-puzzle 7	
Permutaciones a visitar	$16*15*14*13*12*11*10*9 = 518918400$
Número de entradas	$16*15*14*13*12*11*10 = 57657600$

15-puzzle 8	
Permutaciones a visitar	$16*15*14*13*12*11*10*9*8 = 4151347200$
Número de entradas	$16*15*14*13*12*11*10*9 = 518918400$

24-puzzle 4	
Permutaciones a visitar	$25*24*23*22*21 = 6375600$
Número de entradas	$25*24*23*22 = 303600$

24-puzzle 5	
Permutaciones a visitar	$25*24*23*22*21*20 = 127512000$
Número de entradas	$25*24*23*22*21 = 6375600$

24-puzzle 6	
Permutaciones a visitar	$25*24*23*22*21*20*19 = 2422728000$
Número de entradas	$25*24*23*22*21*20 = 127512000$

En las tablas anteriores se presentan los cálculos necesarios por sección del PDB. Particularmente, el PDB 7-8 para el 15-puzzle y el PDB 6-6-6-6 para el 24-puzzle, tienen un espacio de búsqueda demasiado extenso para poder ser almacenados en la memoria sin una técnica más avanzada de la que se está utilizando, por lo que se decidió manejar únicamente las particiones de 5-5-5 y 6-6-3 para el 15-puzzle y las de 5-5-5-5-4 para el 24-puzzle.

El método utilizado para almacenar los valores de los costos hasta el goal del PDB, viene dado por archivos ubicados en la carpeta “pddb”. En ella también se encuentran las divisiones seleccionadas en cada caso, utilizando una función de Hash perfecta, basada en una función biyectiva que incluye combinatorios y matemática discreta. Para la carga de los PDBs en memoria, se utiliza este precálculo, leído desde un archivo para llenar un arreglo en el que cada posición representa un patrón del tablero.

A continuación, se mostrará una serie de experimentos realizados recientemente, utilizando tanto el algoritmo de A* como el algoritmo de IDA*, y que muestran cómo una buena heurística puede ayudar a reducir el tiempo para la solución de problemas complejos. Se observará cómo la heurística de la Distancia Manhattan es mucho menos potente que la heurística de Manhattan con Linear Collision, así como también esta es mucho menos potente que las heurísticas de Pattern Database (PDB). También se apreciará cómo parece haber una tendencia que indica que a medida que un patrón (en la heurística de Pattern Database) sea un subconjunto más grande de las piezas, parece existir una mejora en el tiempo de búsqueda.

Esta conclusión a la que hemos llegado se base principalmente en la interacción del movimiento de varias fichas simultáneamente, mientras Manhattan no toma en consideración ningún tipo de colisión, las heurísticas basadas en PDBs divide el problema en secciones que conviven buscando el goal. Para confirmar esto hicimos una comparación de 2 patrones 5-5-5 con diferente división de fichas (menos efectiva por permitir menos colisión), en las cuáles se observa claramente que las divisiones son determinantes porque de ellas se derivan las posibles interacciones.

Los resultados están basados en las 100 instancias de los casos que provee el profesor Blai Bonet:

Resultados utilizando el algoritmo de A*:

\Heurística Instancia\	Manhattan	Manhattan con linear collision	PDB 5-5-5	PDB 6-6-3	PDB 5-5-5 (menos efectiva)
1	4.49	0.78	0.05	0.03	0.18
2	87.43	7.65	0.12	0.13	5.13
3	23.25	3.8	0.03	0.1	0.77
4	156.41	46.37	1.77	1.45	9.21
5	52.66	12.86	0.26	0.51	1.17
6	10.22	2.89	0.89	0.25	0.91
7	3.8	0.94	0.18	0.08	0.27
8	52.36	10.95	0.29	0.33	3.98
9	20.59	1.78	0.24	0.23	0.78
10	2.25	0.14	0.04	0.04	0.06
11	79.65	19.02	3.17	0.66	2.28
12	27.47	5.19	1.39	0.79	1.52
13	2.15	0.32	0.1	0.16	0.12
14	3.19	0.36	0.12	0.07	0.04
15	159.78	24.29	0.4	1.24	2.25
16	Falla por memoria	44.29	8.41	5.98	8.55
17	1.83	4.25	0.8	1.53	0.79
18	Falla por memoria	86.76	8.7	4.43	39.81

19	10.07	9.68	0.79	1.23	4.23
20	1.95	0.44	0.01	0.08	0.1
21	15.79	3.21	0.28	0.15	0.72
22	38.45	5.32	0.99	1.06	3.78
23	121.48	14.35	3.26	1.24	3.94
24	21.01	1.94	0.29	0.26	0.34
25	26.7	7.41	0.9	0.62	1.54
26	32.72	5	0.52	0.48	1.1
27	81.23	22.41	3.79	1.42	4.35
28	80.42	12.74	2.56	0.46	2.42
29	9.66	1.75	0.21	0.08	0.22
30	17.79	2.3	0.15	0.18	0.37
31	2.18	0.33	0.03	0.07	0.03
32	0.8	0.31	0.01	0.04	0
33	127.49	20.4	4.09	0.77	9.97
34	89.44	12.03	2.23	0.85	1.65
35	16.23	1.69	0.34	0.25	0.2
36	44.42	7.78	1.54	0.27	1.35
37	12.22	2.33	0.25	0.13	0.5
38	80.52	9.67	0.45	0.39	2.32
39	12.72	2.41	0.18	0.22	0.33
40	19.76	7.59	0.4	0.44	0.44
41	17.99	8.46	0.2	0.22	0.88
42	1.82	0.69	0.01	0.05	0.09
43	18.37	4.47	0.9	0.17	0.85
44	10	0.8	0.15	0.13	0.19

45	3.61	0.19	0.05	0.08	0.14
46	7.65	0.46	0.11	0.07	0.16
47	1.03	0.21	0.02	0.05	0.04
48	0.7	0.1	0.02	0.05	0.02
49	Fala por memoria	74.83	4.12	2.92	6.26
50	10.43	9.1	0.4	0.83	0.64
51	9.54	2.1	0.14	0.17	0.09
52	81.03	9.02	0.3	0.49	0.29
53	Falla por memoria	71.15	3.37	2.2	8.35
54	43.55	15.9	0.6	0.78	1.03
55	4.82	0.8	0.04	0.1	0.03
56	241.52	51.55	4.31	2.53	3.99
57	29.27	5.15	0.41	0.68	0.37
58	3.26	0.93	0.13	0.13	0.03
59	Falla por memoria	44.3	2.47	1.57	7.45
60	Falla por memoria	89.7	18.52	4.86	16.13
61	1.91	7.63	1.64	1.25	1.43
62	18.16	2.3	1.15	0.55	2.58
63	98.37	13.71	0.42	0.29	7.18
64	43.11	11.02	0.3	0.31	1.44
65	9.08	1.18	0.1	0.14	0.17
66	Falla por memoria	84.52	4.1	1.49	9.48
67	73.68	17.41	1.01	0.95	1.73

68	41.22	5.52	0.23	0.26	0.7
69	36.48	1.47	4.16	0.31	1.82
70	24.58	10.31	1.46	0.81	1.03
71	5.18	0.97	0.11	0.24	0.14
72	192.74	18.76	1.76	0.56	3.22
73	21.56	1.9	0.16	0.18	0.3
74	1.94	0.38	0.04	0.09	0.27
75	14.1	2.07	1.48	0.32	0.43
76	36.68	9.21	2.56	1.51	0.84
77	11.12	2.63	0.88	0.65	0.68
78	3.86	0.52	0.17	0.18	0.12
79	0.47	0.08	0.02	0.08	0.02
80	25.8	7.66	0.69	0.65	1.86
81	6.24	0.95	0.07	0.21	0.21
82	Falla por memoria	170.34	4.32	2.41	12.18
83	7.58	15.64	0.42	0.73	1.01
84	111.7	19.65	0.34	0.38	1.61
85	12.6	1.74	0.04	0.14	0.15
86	0.66	0.06	0.01	0.08	0.02
87	18.65	2.5	0.06	0.12	0.42
88	Falla por memoria	184.76	21.17	9.03	90.82
89	20.94	20.53	2.8	2.65	9.28
90	5.05	0.54	0.29	0.16	0.21
91	212.3	21.15	0.86	0.27	1.71

92	206.33	21.74	2.04	0.7	2.01
93	22.46	2.17	0.17	0.23	0.19
94	0.43	0.09	0.02	0.11	0.02
95	5.31	1.48	0.02	0.12	0.14
96	8.39	0.46	0.13	0.15	0.38
97	1.39	0.23	0.02	0.09	0.06
98	47.38	21.41	0.74	0.64	1.29
99	50.11	9.11	1.23	0.6	1.01
100	23.41	8.38	1.03	0.71	1.42
Promedio de los 100 casos	Hubo casos que no se pudo determinar su tiempo	15.1582	1.4472	0.7713	3.2955

Resultados utilizando el algoritmo de IDA*:

\Heurística Instancia\	Manhattan	Manhattan con linear collision	PDB 5-5-5	PDB 6-6-3	PDB 5-5-5 (menos efectiva)
1	0.96	0.32	0.02	0.03	0.05
2	14.64	5.83	0.05	0.05	1.11
3	1.48	1.3	0.02	0.08	0.12
4	43.8	34.95	0.67	1.73	5.67
5	5.69	3.03	0.05	0.05	0.09
6	0.79	0.83	0.25	0.11	0.25
7	0.91	0.26	0.03	0.02	0.05
8	8.46	7.95	0.05	0.06	0.67
9	2.49	0.481	0.15	0.11	0.25

10	0.32	0.05	0	0.02	0.02
11	14.2	32.09	0.7	0.48	0.79
12	7.69	1.94	0.61	0.84	0.64
13	0.04	0.04	0	0.01	0.01
14	1.03	0.46	0.07	0.03	0.02
15	39.08	3.82	0.3	0.46	0.55
16	20.1	14.28	2.46	3.01	2.38
17	1.18	0.57	0.08	0.07	0.03
18	28.05	49.97	1.83	3.03	4.06
19	1.64	1.33	0.04	0.04	0.2
20	0.09	0.24	0.01	0.01	0.01
21	2.84	2.45	0.08	0.18	0.28
22	17.3	2.98	0.58	0.98	2.89
23	25.54	7.13	0.85	0.91	0.88
24	1	0.51	0.01	0.01	0.02
25	3.1	3.93	0.29	0.54	1
26	6.19	4.44	0.33	0.35	0.39
27	18.51	12.15	1.32	1.42	3.05
28	13	6.67	1.31	0.09	1.93
29	0.59	0.48	0.02	0.02	0.04
30	3.99	1.85	0.1	0.12	0.29
31	0.14	0.16	0.01	0.01	0
32	0.11	0.19	0.01	0.01	0.01
33	23.85	31.24	2.26	2..06	5.22
34	35.25	9.46	0.48	0.3	0.73
35	3.34	0.85	0.16	0.14	0.04

36	2.72	1.8	0.17	0.13	0.38
37	4.6	0.63	0.06	0.05	0.26
38	21.22	6.02	0.16	0.33	1.03
39	1.46	1.21	0.17	0.11	0.08
40	5.2	7.15	0.19	0.09	0.14
41	4.85	6.43	0.07	0.05	0.34
42	0.08	0.03	0	0.01	0.01
43	3.43	1.93	0.23	0.08	0.25
44	4.55	0.62	0.08	0.25	0.08
45	0.73	0.35	0.06	0.05	0.15
46	1.27	0.4	0.04	0.01	0.11
47	0.11	0.06	0.01	0	0.01
48	0.07	0.02	0.01	0.01	0.01
49	122.25	68.11	2.06	2.62	1.53
50	3.39	0.7	0.05	0.02	0.17
51	2.35	1.29	0.05	0.06	0.04
52	23.21	6.72	0.24	0.31	0.41
53	30.8	29.74	0.81	0.88	2.21
54	15.96	7.4	0.15	0.23	0.33
55	0.05	0.02	0.01	0.01	0.01
56	78.2	38.47	2.64	1.65	1.83
57	0.46	0.69	0.01	0	0.03
58	0.83	0.77	0.1	0.06	0.04
59	88.31	24.36	1.54	1.14	2.92
60	219.74	49.81	5.5	2.12	4.49

61	0.47	0.17	0.01	0.01	0.04
62	3.96	5.21	0.82	0.56	1.28
63	70.01	3.1	0.19	0.19	2.87
64	16.88	5.69	0.14	0.15	0.57
65	1.43	0.62	0.08	0.02	0.05
66	106.78	65.97	1.42	0.89	6.27
67	11.6	12.78	0.38	0.18	0.29
68	11.43	1.99	0.13	0.16	0.26
69	9.78	2.27	0.86	0.24	1.13
70	10.61	10.18	0.71	0.6	0.83
71	0.34	0.08	0.01	0	0.02
72	116.32	21.28	0.64	0.46	1.53
73	0.11	0.13	0.01	0	0.01
74	0.11	0.07	0.01	0	0.04
75	3.67	2.45	0.52	0.19	0.26
76	8.66	24.38	1.89	0.35	0.22
77	1.21	1.2	0.1	0.05	0.26
78	0.79	0.58	0.07	0.06	0.16
79	0.02	0.58	0	0.01	0
80	10.08	10.13	0.38	0.58	1.16
81	1.38	0.45	0.01	0.01	0.11
82	324.24	115.63	2.44	0.94	6.53
83	3.48	0.29	0.05	0.01	0.11
84	9.53	18.26	0.12	0.13	0.3
85	0.06	0.07	0.01	0	0.01
86	0.16	0.03	0.01	0.01	0.01

87	4.75	2.8	0.04	0.01	0.24
88	500.4	121.08	5.96	2.97	43.38
89	9.71	4.61	0.87	0.32	0.6
90	0.89	0.29	0.04	0.04	0.11
91	84.74	41.88	0.42	0.31	2.18
92	79.21	23.19	0.68	0.66	0.89
93	0.62	0.5	0	0	0.04
94	0.03	0.04	0.01	0.01	0.01
95	0.62	1.13	0.02	0.01	0.05
96	0.82	0.2	0.06	0.06	0.2
97	0.25	0.14	0	0	0.02
98	11.49	9.57	0.75	0.82	1.01
99	9.42	6.49	0.5	0.08	0.13
100	3.52	2.83	0.22	0.21	1.21
Promedio de los 100 casos	24.1281	10.1674	0.5018	0.3689	1.3293

Resultados de los tiempos de IDA* para el 24-puzzle:

\ Heurística Casos\	Instancia	PDB 5-5-5-5-4
1	6 5 3 4 9 1 8 2 18 13 0 12 11 23 14 10 15 7 16 19 20 21 17 22 24	0
2	0 5 6 3 4 10 2 1 7 8 16 15 13 18 12 20 11 14 19 9 22 17 21 23 24	0.09
3	10 1 7 4 9 2 6 15 3 13 11 5 12 17 0 20 16 18 14 8 21 22 19 23 24	0.34
4	1 6 2 4 9 10 5 11 8 3 20 15 19 13 18 7 12 23 16 17 21 22 0 24 14	0.01
5	1 0 4 13 9 10 3 6 19 14 21 16 5 2 8 11 15 12 23 18 20 22 17 7 24	1.32
6	10 2 7 9 3 13 15 6 4 14 16 0 5 1 8 11 12 18 24 19 20 21 17 22 23	9.41
7	5 0 2 8 4 10 12 7 9 14 15 1 6 3 13 20 11 17 18 19 21 16 22 23 24	0.37
8	1 7 3 8 4 0 6 5 13 9 21 11 12 14 18 10 2 15 17 16 20 22 23 24 19	0.43
9	1 13 7 9 2 11 6 3 4 14 16 15 12 19 5 20 17 18 8 0 21 22 10 23 24	21.68
10	12 0 9 7 8 1 5 6 3 4 11 10 17 13 14 15 2 16 18 19 20 21 22 23 24	0.81
11	5 1 7 8 3 10 19 6 2 0 15 23 11 4 9 20 16 12 17 14 21 22 18 13 24	13.24
12	2 13 4 0 9 1 6 8 16 12 5 15 3 7 24 11 10 21 18 14 20 22 19 23 17	10.19
13	1 10 2 3 4 5 12 15 9 14 17 0 7 6 19 16 11 22 8 13 20 21 23 18 24	0.06
14	6 3 12 8 4 1 7 11 2 9 5 23 13 0 14 15 20 21 18 19 16 10 22 17 24	1.45
15	5 3 4 8 9 6 1 7 2 14 21 15 10 12 13 0 17 16 18 19 11 20 22 23 24	16.01
16	5 1 8 2 4 6 7 13 16 9 10 11 17 3 19 15 14 24 22 23 20 0 21 12 18	0.6

17	5 1 2 3 4 11 10 7 8 9 15 6 13 22 14 23 16 0 18 19 20 21 12 17 24	0.2
18	11 1 2 8 4 16 3 6 9 14 10 17 12 13 19 20 7 5 18 0 21 15 22 23 24	0.15
19	6 5 10 3 4 11 7 12 2 13 15 16 18 9 8 17 1 0 19 14 20 21 22 23 24	0.44
20	6 8 4 14 9 1 5 7 3 2 15 10 18 24 13 0 11 17 12 23 20 16 21 22 19	7.79
21	5 3 12 8 13 2 11 4 1 9 10 16 6 7 18 20 15 17 19 14 21 0 22 23 24	2.56
22	5 1 2 9 3 0 10 11 7 4 23 6 13 8 14 15 16 22 17 24 20 21 12 19 18	119.81
23	1 2 8 13 9 3 10 12 19 4 15 5 7 6 14 0 21 11 17 18 16 20 22 23 24	0.05
24	11 1 6 3 4 5 16 0 2 9 10 7 12 8 14 15 21 17 13 19 20 23 22 18 24	62.1
25	6 3 13 9 7 1 2 8 11 4 5 10 20 0 14 21 16 17 12 18 15 22 23 24 19	1.04

Otros resultados:

- La cantidad de nodos generados por manhattan es en promedio 4×10^6 por segundo.
- La cantidad de nodos generados por manhattan pairwise es en promedio 1.67×10^6 por segundo.
- La cantidad de nodos generados por PDB 5-5-5 es en promedio 1.39×10^6 por segundo.
- La cantidad de nodos generados por PDB 6-6-3 es en promedio 1.48×10^6 por segundo.
- La cantidad de nodos generados por PDB 5-5-5-5-4 es en promedio 3.12×10^6 por segundo.
- Los thresholds de todas las corridas de IDA* tienen un aumento constante en todas las corridas de 2 movimientos (costo).
- El código posee un readme.txt en el que se explican detalles para la corrida.
- En la carpeta PDB se encuentran los costos asociados para cada pdb, del mismo modo que la división para cada base de datos.
- En la carpeta resultados se encuentra una versión extendida por caso de cada solución de cada instancia con cada heurística.