



Tarea 5

1 Pruebas

En la siguiente tabla, se detallan los tiempos de ejecución de ambos algoritmos, **fuerzabruta** y **mejorado**, para los tests entregados. Como se puede notar, no se incluyen los resultados de los tests de 50 proposiciones para el algoritmo **fuerzabruta**, debido a que no fue capaz de terminar ninguna ejecución en una cantidad razonable de tiempo. Las razones se discuten posteriormente.

# test	fuerzabruta (s)	mejorado (s)
1	1.478164	0.003737
2	1.541634	0.004029
3	0.126838	0.011579
4	1.100794	0.003986
5	3.540519	0.006219
6	0.771406	0.000953
7	2.778702	0.003771
8	2.755018	0.008069
9	2.242883	0.006876
10	1.034398	0.008547
11	0.15608	0.001644
12	0.602816	0.001112
13	3.616978	0.000864
14	0.050678	0.00141
15	3.2572	0.000908
16	1.983858	0.000897
17	0.552265	0.006315
18	2.208769	0.008461
19	2.750886	0.005064
20	0.674245	0.001043
AVG	1.661207	0.004274
MDN	1.509899	0.003879

Figure 1: Tiempos de ejecución para tests de 20 proposiciones (insatisfacibles)

# test	mejorado (s)	# test	mejorado (s)
1	1.394208	1	54.115868
2	33.4989	2	32.340251
3	0.081594	3	22.96283
4	23.804611	4	35.996101
5	6.294908	5	20.785146
6	36.309804	6	38.294041
7	16.575414	7	44.600387
8	14.90874	8	39.513315
9	19.801974	9	116.971578
10	4.658373	10	13.990002
AVG	15.7328526	AVG	41.9569519
MDN	15.742077	MDN	37.145071

(a) satisfacibles (b) insatisfacibles

Figure 2: Tiempos de ejecución para tests de 50 proposiciones

Para los primeros tests (de 20 proposiciones), ambos algoritmos tienen tiempos bastante reducidos. Para el caso del algoritmo de fuerza bruta, en el cual se prueban todas las valuaciones posibles hasta encontrar la solución (que en estos tests siempre hay, debido a que son todos satisfacibles), la cantidad máxima de valuaciones que debe probar es de $2^{20} = 1048576$, que si bien es un número superior al millón, un computador moderno no debería tener problemas en manejarlo. Además, dado que los tests son satisfacibles, el número corresponde al peor caso absoluto, y es muy probable que las soluciones se hayan encontrado bastante antes de acercarse al número. Para el algoritmo mejorado, los tiempos son ínfimos, llegando a ser en promedio casi 400 veces más rápido.

Para los tests de 50 proposiciones (y como se puede notar a partir de la tabla), le es imposible al algoritmo de fuerza bruta encontrar una solución en un tiempo razonable. Ahora, la cantidad de valuaciones posibles a probar escala a $2^{50} = 1125899906842624$ (!!!). Además, dado que los tests de la última tabla son insatisfacibles, ese es exactamente el número de soluciones que hay que probar. Este número llega a ser inmanejable, por lo que es prácticamente imposible resolver problemas de este orden con este algoritmo. Esto se reflejó en la práctica, donde el algoritmo fue incapaz de llegar a un resultado luego de varios minutos, lo cual, luego de considerar la gran cantidad de valuaciones, no es muy sorprendente.

En el caso del algoritmo mejorado, para los tests de 50 proposiciones, se logró implementar un SAT solver capaz de resolver los tests en tiempos razonables. Como se puede apreciar de las tablas, los tiempos de los tests insatisfacibles tienden a ser mayores que los tiempos para los tests satisfacibles, lo cual hace bastante sentido, debido a que en estos últimos, basta encontrar una solución para llegar a una respuesta. Por otro lado, en los tests insatisfacibles, hay que agotar todas las respuestas posibles, sin embargo, las iteraciones que se hacen son mucho menores a 2^{50} , gracias a las podas/heurísticas implementadas (en base a DPLL).

Personalmente, luego de bastante ensayo y error, me sorprendió haber llegado a un algoritmo que lograra resolver tests de 50 proposiciones, y también, al hacer el análisis, me sorprendió la cantidad de posibles valuaciones que tiene que probar el algoritmo mejorado, en comparación al de fuerza bruta. Esto último llama la atención particularmente para los tests insatisfacibles, debido a que se debe llegar a un estado en el que se concluya con seguridad que no hay ninguna valuación que satisfaga a la fórmula, lo cual se realiza probando una fracción muy pequeña del espacio total de valuaciones.