

COSC-4117EL: Assignment 1 Report

Group Number: 2

Group Member:

NAME	STUDENT#	EMAIL	CONTRIBUTION
Haoliang Sheng	0441916	hsheng@laurentian.ca	60%
Zihao Zhou	0429993	zzhou3@laurentian.ca	20%
Jiazhou Ye	0426609	Jye1@laurentian.ca	20%

1. Abstract

In this study, we developed a Python program to solve binary puzzles using constraint satisfaction algorithms, specifically focusing on three heuristics: Random Variable (H1), Most Constrained Variable (H2), and Most Constraining Variable (H3). We further explored the impact of employing the AC3 constraint propagation method within these algorithms. By utilizing puzzles sourced from the [Binary Puzzle](#) website, our evaluations spanned various sizes and difficulty levels. Performance metrics included average solving time, search attempts, and completion rates. Our primary findings revealed the robustness of the H2 heuristic across all scenarios and the consistent performance enhancements offered by AC3 in terms of search attempts.

2. Introduction

Binary puzzles are a captivating subset of constraint satisfaction problems. Presented as square grids of size $n \times n$, the cells within these puzzles can either be filled with a 0, a 1, or left blank. The solver's objective is to populate the blank cells in a manner that adheres to a set of predefined rules:

- a) Each cell must contain a number: either 0 or 1.
- b) Three consecutive ones or zeros cannot appear in any row or column.
- c) Every row and column must have an equal number of ones and zeros.
- d) No two rows or columns can be identical.

In our endeavor to assess the efficiency of different heuristics and constraint propagation methods, puzzles were sourced from the Binary Puzzle website. For each puzzle size (ranging from 6x6 to 14x14) and difficulty level (easy, medium, hard, very hard), five distinct puzzles were transformed into a .txt format.

3. Methodology

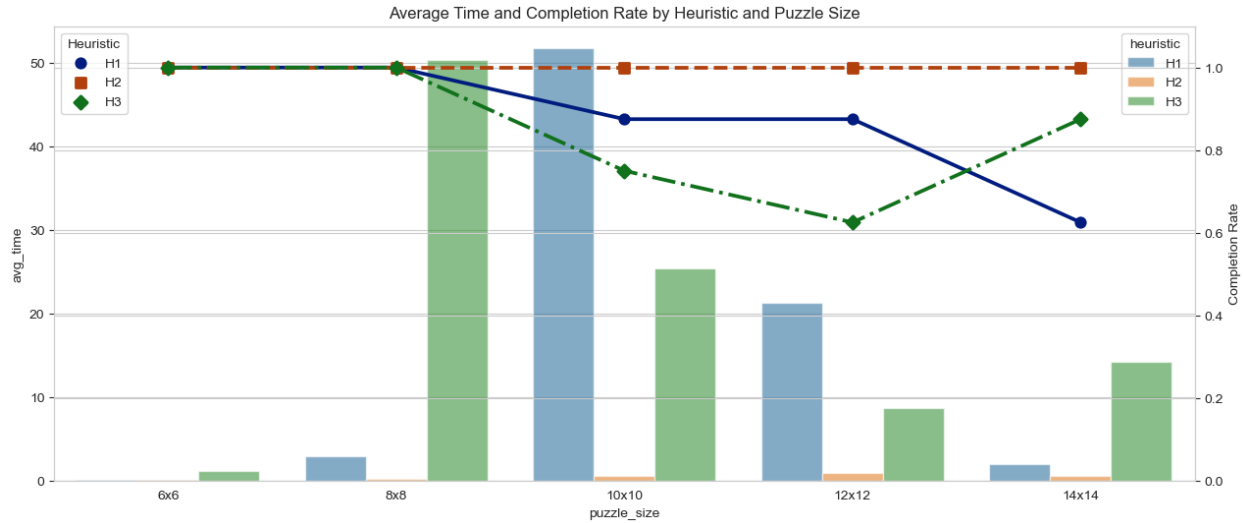
Initially, the approach involved averaging the performance metrics across the five different puzzles for a specific combination of size, difficulty, heuristic, and AC3 usage. However, this method posed challenges due to inconsistencies in completion times, leading to potential unfairness in averaged results.

Subsequently, a revised approach was adopted: a single puzzle, under a specified set of conditions, was run ten times. The average time and attempts across these ten runs were then computed. If a particular set of conditions resulted in solving times exceeding 5 minutes, the case was labeled as "time out" and deemed incomplete.

4. Experimental Results

Impact of Heuristics on Different Puzzle Sizes with Completion Overlay

	heuristic	puzzle_size	avg_time	avg_attempts	is_complete
0	H1	10x10	51.827795	30825.428571	0.875
1	H1	12x12	21.330801	8886.714286	0.875
2	H1	14x14	1.998486	327.800000	0.625
3	H1	6x6	0.066898	159.625000	1.000
4	H1	8x8	2.924404	4267.125000	1.000
5	H2	10x10	0.531899	277.875000	1.000
6	H2	12x12	0.874933	204.750000	1.000
7	H2	14x14	0.568549	57.625000	1.000
8	H2	6x6	0.041968	55.375000	1.000
9	H2	8x8	0.147971	105.125000	1.000
10	H3	10x10	25.402272	201.333333	0.750
11	H3	12x12	8.708907	49.400000	0.625
12	H3	14x14	14.167051	53.714286	0.875
13	H3	6x6	1.108602	159.625000	1.000
14	H3	8x8	50.378604	4267.125000	1.000



The plot above displays the average time taken by different heuristics across various puzzle sizes, with the completion rate overlaid:

- Bars represent the average time taken.
- Lines represent the completion rate.

Time Observations:

H1:

Average time peaks at 51.83 seconds for 10x10 puzzles, then decreases significantly to 21.33 seconds for 12x12 puzzles, and further drops to 1.998 seconds for 14x14 puzzles.

The completion rate for H1 remains at 0.875 for both 10x10 and 12x12 puzzles but decreases to 0.625 for 14x14 puzzles, indicating that some puzzles might not have been completed within the time limit.

H2:

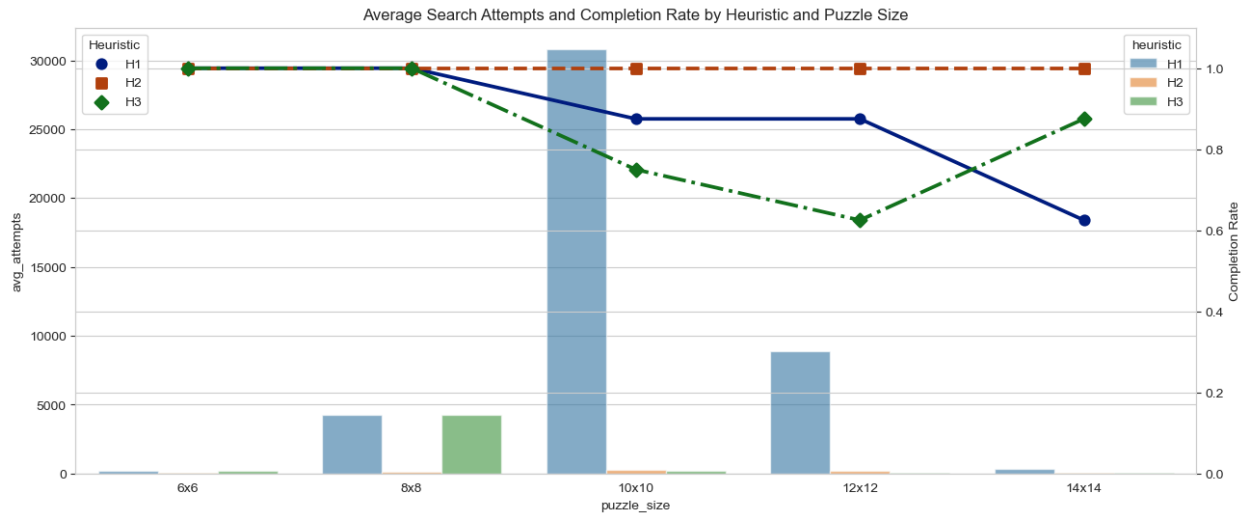
Maintains a consistently low average time across all puzzle sizes, with the highest being 0.8749 seconds for 12x12 puzzles.

Achieves a 100% completion rate across all puzzle sizes, indicating its robustness.

H3:

Shows a spike in average time for 8x8 puzzles at 50.38 seconds, then decreases to 25.40 seconds for 10x10 puzzles. There's another increase to 14.17 seconds for 14x14 puzzles.

The completion rate remains high for most puzzle sizes but dips to 0.625 for 12x12 puzzles.



The plot above displays the average number of search attempts made by different heuristics across various puzzle sizes, with the completion rate overlaid:

- Bars represent the average number of search attempts.
- Lines represent the completion rate.

Attempts Observations:

H1:

The average number of search attempts peaks at 30,825 for 10x10 puzzles, decreases to 8,886 for 12x12 puzzles, and further drops to 327.8 for 14x14 puzzles.

The completion rate for H1 remains at 0.875 for both 10x10 and 12x12 puzzles but decreases to 0.625 for 14x14 puzzles, suggesting that some puzzles might not have been completed within the time limit.

H2:

Maintains a consistently low number of search attempts across all puzzle sizes, with the highest being 277.875 for 10x10 puzzles.

Achieves a 100% completion rate across all puzzle sizes, reinforcing its efficiency.

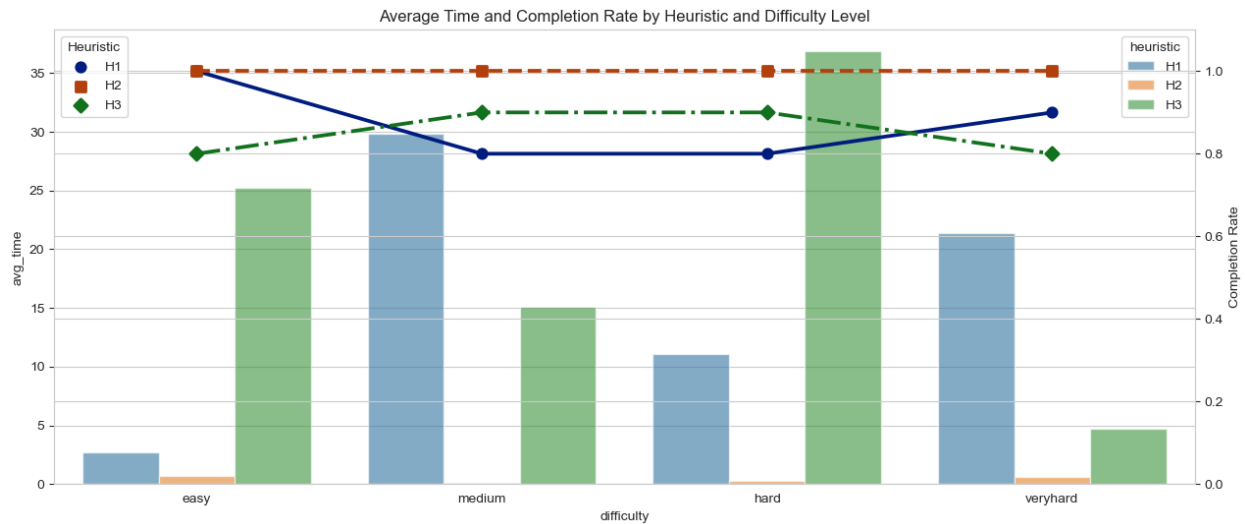
H3:

Like H1, H3 also shows a spike in search attempts for 8x8 puzzles with 4,267.125 attempts. This decreases to 201.333 attempts for 10x10 puzzles and slightly increases to 53.714 for 14x14 puzzles.

The completion rate remains high for most puzzle sizes but dips to 0.625 for 12x12 puzzles.

Impact of Heuristics on Different Puzzle Difficulties with Completion Overlay

	heuristic	difficulty	avg_time	avg_attempts	is_complete
0	H1	easy	2.668019	614.500000	1.0
1	H1	hard	11.062706	7026.000000	0.8
2	H1	medium	29.828063	15581.000000	0.8
3	H1	veryhard	21.358519	14226.333333	0.9
4	H2	easy	0.691424	115.100000	1.0
5	H2	hard	0.259002	88.400000	1.0
6	H2	medium	0.142124	51.100000	1.0
7	H2	veryhard	0.639705	306.000000	1.0
8	H3	easy	25.210654	209.500000	0.8
9	H3	hard	36.836982	3444.000000	0.9
10	H3	medium	15.137281	240.777778	0.9
11	H3	veryhard	4.696447	300.750000	0.8



The plot above displays the average time taken by different heuristics across various difficulty levels, with the completion rate overlaid:

- Bars represent the average time taken.
- Lines represent the completion rate.

Time Observations:

H1:

Performs most efficiently on 'easy' puzzles with an average time of 2.668 seconds.

For 'medium' puzzles, the average time taken by H1 spikes to 29.83 seconds, which is even higher than the 21.36 seconds taken for 'very hard' puzzles.

The completion rate for H1 is perfect for 'easy' puzzles but drops to 0.8 for both 'hard' and 'medium' puzzles. Interestingly, the completion rate rises slightly to 0.9 for 'very hard' puzzles.

H2:

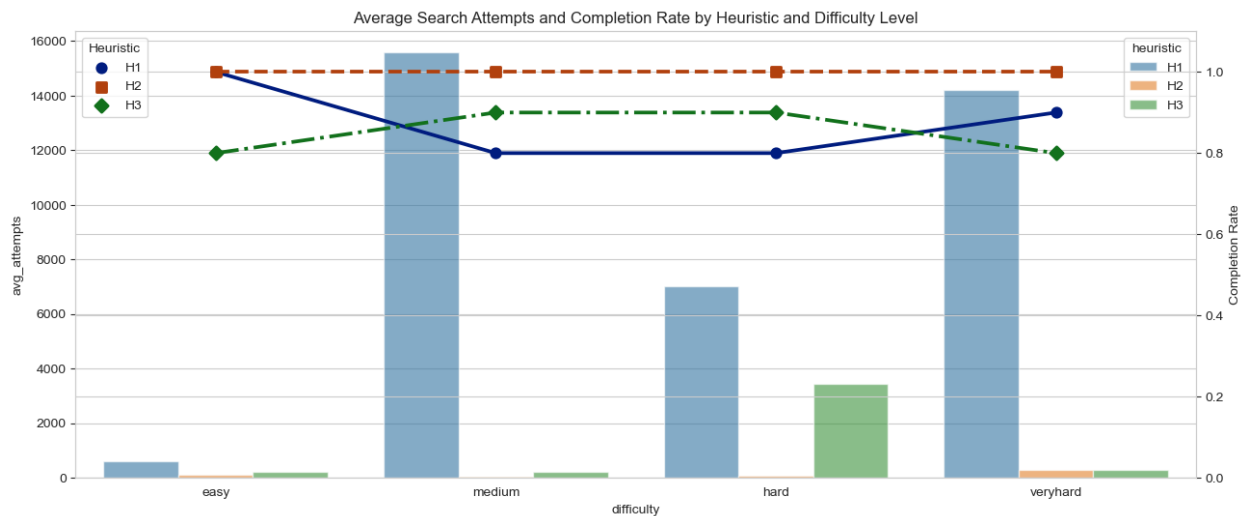
Maintains a consistently low average time across all difficulty levels. While the time is lowest for 'medium' puzzles at 0.1421 seconds, there's a slight increase to 0.6397 seconds for 'very hard' puzzles.

Achieves a 100% completion rate across all difficulty levels.

H3:

The average time taken by H3 is highest for 'hard' puzzles at 36.84 seconds, whereas for 'very hard' puzzles, the time surprisingly drops to 4.696 seconds.

The completion rate remains at 0.9 for 'medium' and 'hard' puzzles but drops to 0.8 for both 'easy' and 'very hard' puzzles.



The plot above displays the average number of search attempts made by different heuristics across various difficulty levels, with the completion rate overlaid:

- Bars represent the average number of search attempts.
- Lines represent the completion rate.

Attempts Observations:

H1:

The number of search attempts is moderate for 'easy' puzzles at 614.5 but jumps significantly to 15,581 for 'medium' puzzles. For 'very hard' puzzles, it slightly decreases to 14,226.33.

The completion rate is perfect for 'easy' puzzles but drops to 0.8 for both 'hard' and 'medium' puzzles. However, it rises slightly to 0.9 for 'very hard' puzzles.

H2:

Maintains a consistently low number of search attempts across all difficulty levels, with the highest being 306 for 'very hard' puzzles.

Achieves a 100% completion rate across all difficulty levels, reinforcing its efficiency and robustness.

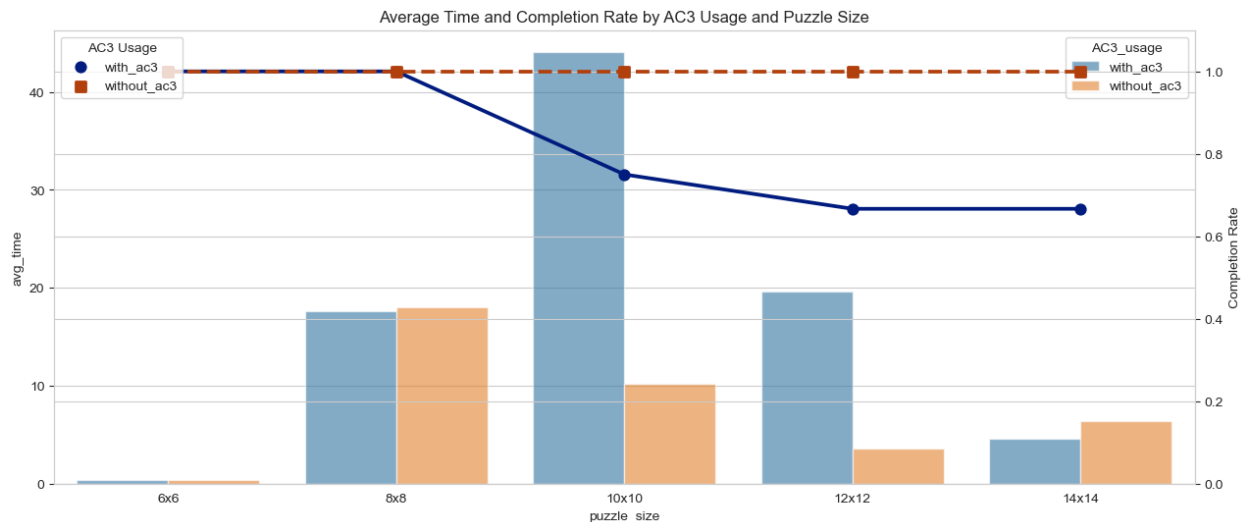
H3:

The number of search attempts is highest for 'hard' puzzles at 3,444. For 'very hard' puzzles, the attempts drop to 300.75.

The completion rate remains at 0.9 for 'medium' and 'hard' puzzles but drops to 0.8 for both 'easy' and 'very hard' puzzles.

Impact of AC3 Usage on Different Puzzle Sizes with Completion Overlay

	AC3_usage	puzzle_size	avg_time	avg_attempts	is_complete
0	with_ac3	10x10	44.107786	5892.777778	0.750000
1	with_ac3	12x12	19.575666	1485.000000	0.666667
2	with_ac3	14x14	4.561190	66.750000	0.666667
3	with_ac3	6x6	0.411334	71.083333	1.000000
4	with_ac3	8x8	17.630318	1352.750000	1.000000
5	without_ac3	10x10	10.207776	13847.833333	1.000000
6	without_ac3	12x12	3.604523	4351.000000	1.000000
7	without_ac3	14x14	6.435055	161.833333	1.000000
8	without_ac3	6x6	0.400311	178.666667	1.000000
9	without_ac3	8x8	18.003668	4406.833333	1.000000



The plot above displays the impact of AC3 usage on the average time taken across various puzzle sizes, with the completion rate overlaid:

- Bars represent the average time taken.
- Lines represent the completion rate, differentiated by AC3 usage (Yes or No).

Time Observations:

Without AC3:

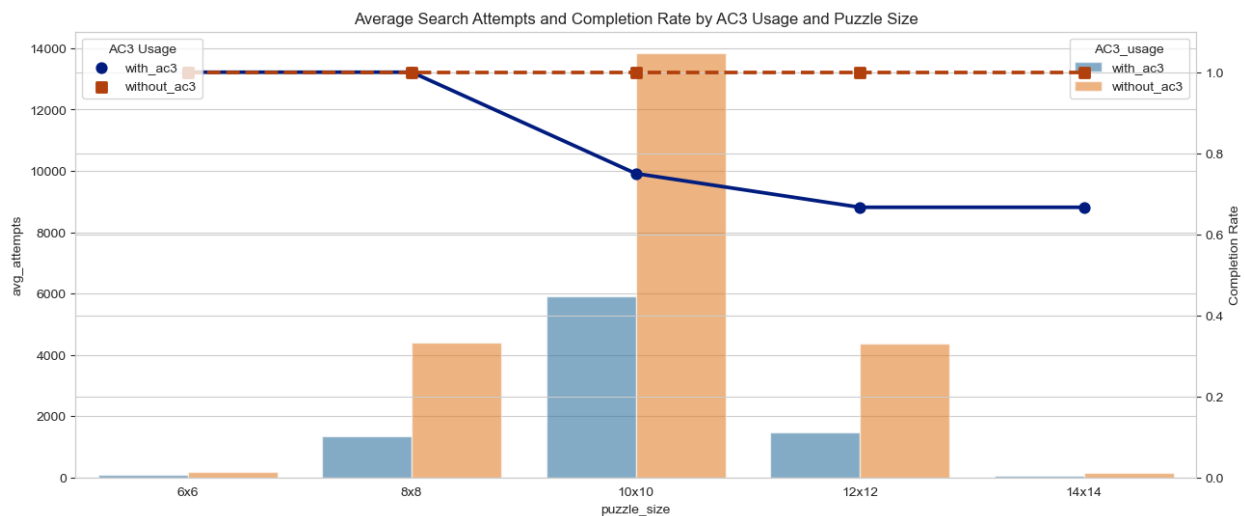
The average time peaks at 10.21 seconds for 10x10 puzzles, which is surprisingly lower than the time taken for 8x8 puzzles at 18.00 seconds.

The completion rate without AC3 remains at 100% across all puzzle sizes.

With AC3:

The average time is highest for 10x10 puzzles at 44.11 seconds, which then decreases for larger puzzles such as 12x12 and 14x14, with times of 19.58 seconds and 4.56 seconds respectively.

The completion rate remains at 100% for smaller puzzles like 6x6 and 8x8. However, it drops to 0.75 for 10x10 puzzles and further decreases to 0.67 for both 12x12 and 14x14 puzzles.



The plot above displays the impact of AC3 usage on the average number of search attempts across various puzzle sizes, with the completion rate overlaid:

- Bars represent the average number of search attempts.
- Lines represent the completion rate, differentiated by AC3 usage (Yes or No).

Attempts Observations:

Without AC3:

The average number of search attempts is notably high for 10x10 puzzles with 13,847.83 attempts. For 8x8 puzzles, it jumps to 4,406.83 attempts.

The completion rate without AC3 remains at 100% across all puzzle sizes.

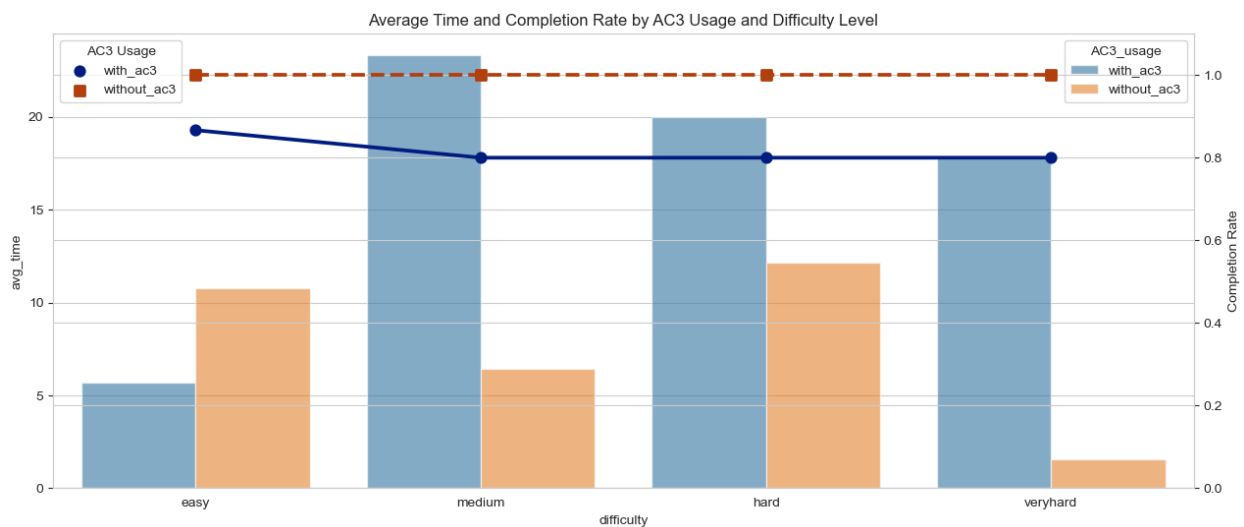
With AC3:

Puzzles solved using AC3 generally require fewer search attempts, with 5,892.78 attempts for 10x10 puzzles and 1,485 attempts for 12x12 puzzles.

The completion rate remains at 100% for smaller puzzles like 6x6 and 8x8. However, it drops to 0.75 for 10x10 puzzles and further decreases to 0.67 for both 12x12 and 14x14 puzzles.

Impact of AC3 Usage on Different Puzzle Difficulties with Completion Overlay

	AC3_usage	difficulty	avg_time	avg_attempts	is_complete
0	with_ac3	easy	5.661409	159.538462	0.866667
1	with_ac3	hard	20.008167	1631.500000	0.800000
2	with_ac3	medium	23.306431	2801.166667	0.800000
3	with_ac3	veryhard	17.765938	2272.416667	0.800000
4	without_ac3	easy	10.778756	459.866667	1.000000
5	without_ac3	hard	12.168434	4567.333333	1.000000
6	without_ac3	medium	6.440274	6247.466667	1.000000
7	without_ac3	veryhard	1.533603	7082.266667	1.000000



The plot above displays the impact of AC3 usage on the average time taken across various difficulty levels, with the completion rate overlaid:

- Bars represent the average time taken.
- Lines represent the completion rate, differentiated by AC3 usage (Yes or No).

Time Observations:

Without AC3:

The average time taken for 'easy' puzzles is 10.78 seconds. Surprisingly, this decreases for 'hard' and 'medium' puzzles to 12.17 seconds and 6.44 seconds, respectively, and further drops to a mere 1.53 seconds for 'very hard' puzzles.

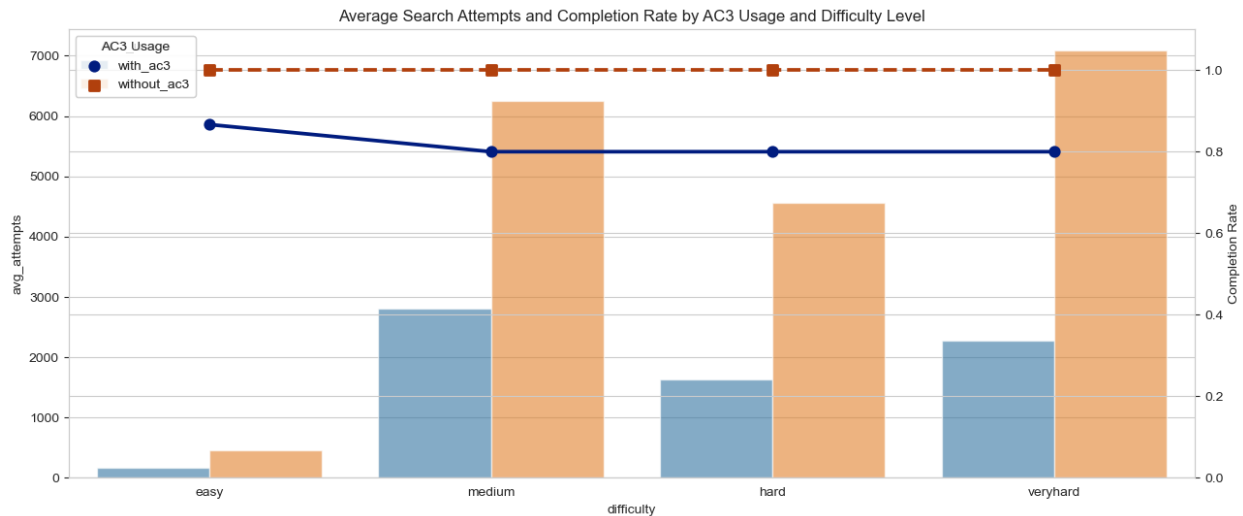
The completion rate without AC3 remains at 100% across all puzzle difficulties.

With AC3:

The average time peaks at 23.31 seconds for 'medium' puzzles. While 'hard' puzzles take 20.01 seconds, the time decreases to 17.77 seconds for 'very hard' puzzles.

In general, puzzles solved using AC3 tend to complete slower across most difficulty levels compared to those solved without AC3.

The completion rate for puzzles solved using AC3 is 0.8 for 'hard', 'medium', and 'very hard' puzzles, and 0.8667 for 'easy' puzzles, which is indeed lower than without using AC3.



The plot above displays the impact of AC3 usage on the average number of search attempts across various difficulty levels, with the completion rate overlaid:

- Bars represent the average number of search attempts.
- Lines represent the completion rate, differentiated by AC3 usage (Yes or No).

Attempts Observations:

Without AC3:

The average number of search attempts is 459.87 for 'easy' puzzles. This number jumps significantly to 4,567.33 for 'hard' puzzles and further increases to 6,247.47 for 'medium' puzzles and 7,082.27 for 'very hard' puzzles.

The completion rate without AC3 remains at 100% across all puzzle difficulties.

With AC3:

The average number of search attempts is 159.54 for 'easy' puzzles. This increases to 1,631.5 for 'hard' puzzles, 2,801.17 for 'medium' puzzles, and 2,272.42 for 'very hard' puzzles.

In general, puzzles solved using AC3 require more search attempts for 'hard', 'medium', and 'very hard' puzzles compared to 'easy' puzzles, but fewer attempts than those solved without AC3 for the same difficulties.

The completion rate for puzzles solved using AC3 is 0.8 for 'hard', 'medium', and 'very hard' puzzles, and 0.8667 for 'easy' puzzles. This rate is lower than without using AC3.

5. Discussion

Heuristic Efficiency:

H2 consistently emerged as the most efficient heuristic. It achieved a 100% completion rate across all scenarios, maintaining low average times and search attempts. This positions it as a

highly reliable choice for solving puzzles. Conversely, H1 and H3 were generally effective but showed diminished performance, especially for larger puzzles. Both heuristics displayed increased times and search attempts as the puzzle size expanded.

Impact of AC3:

Incorporating AC3 into the puzzle-solving process yielded mixed results. Although the use of AC3 consistently led to fewer search attempts, it didn't always translate to quicker solving times. The efficiency regarding time was influenced by the specific puzzle size and difficulty.

General Observations:

With an increase in puzzle size, the challenges inherently magnified, as seen from the declining completion rates for 12x12 and 14x14 puzzles. 'Very Hard' puzzles posed the most significant challenge both in terms of time and search attempts. Interestingly, 'Very Hard' puzzles solved without AC3 demonstrated surprisingly quick completion times, with the average time for such puzzles being a mere 1.53 seconds.

In conclusion, while the heuristic choice plays a pivotal role in solving efficiency, the incorporation of AC3 displayed varied outcomes. Although AC3 consistently minimized the number of search attempts, its effect on solving times wasn't consistent. This analysis underscores the importance of choosing an optimal heuristic, such as H2, and delving into the intricate impacts of utilizing AC3, especially when dealing with larger and intricate puzzles.

6. Appendix: Experimental Results Table

The table below presents the compiled results from the experiments:

REPORT: ASSIGNMENT 1

COURSE: COSC-4117EL-01: Artificial Intelligence

GROUP: 2

uzzle_size	difficulty	heuristic	AC3_usage	avg_time	g_attempts	completed
6x6	easy	H1	with_ac3	0.045558	30	TRUE
6x6	easy	H1	without_a	0.001459	58	TRUE
6x6	medium	H1	with_ac3	0.120849	75	TRUE
6x6	medium	H1	without_a	0.00293	230	TRUE
6x6	hard	H1	with_ac3	0.147917	86	TRUE
6x6	hard	H1	without_a	0.002979	256	TRUE
6x6	veryhard	H1	with_ac3	0.20936	160	TRUE
6x6	veryhard	H1	without_a	0.004133	382	TRUE
6x6	easy	H2	with_ac3	0.035637	24	TRUE
6x6	easy	H2	without_a	0.003949	44	TRUE
6x6	medium	H2	with_ac3	0.084889	28	TRUE
6x6	medium	H2	without_a	0.00434	54	TRUE
6x6	hard	H2	with_ac3	0.074746	40	TRUE
6x6	hard	H2	without_a	0.008995	78	TRUE
6x6	veryhard	H2	with_ac3	0.113762	59	TRUE
6x6	veryhard	H2	without_a	0.009425	116	TRUE
6x6	easy	H3	with_ac3	0.224807	30	TRUE
6x6	easy	H3	without_a	0.182462	58	TRUE
6x6	medium	H3	with_ac3	0.943713	75	TRUE
6x6	medium	H3	without_a	1.24543	230	TRUE
6x6	hard	H3	with_ac3	1.483361	86	TRUE
6x6	hard	H3	without_a	1.993369	256	TRUE
6x6	veryhard	H3	with_ac3	1.451409	160	TRUE
6x6	veryhard	H3	without_a	1.344264	382	TRUE
8x8	easy	H1	with_ac3	0.512307	102	TRUE
8x8	easy	H1	without_a	0.01296	326	TRUE
8x8	medium	H1	with_ac3	1.717276	326	TRUE
8x8	medium	H1	without_a	0.022828	1032	TRUE
8x8	hard	H1	with_ac3	18.96606	6980	TRUE
8x8	hard	H1	without_a	0.307464	23530	TRUE
8x8	veryhard	H1	with_ac3	1.836226	567	TRUE
8x8	veryhard	H1	without_a	0.020111	1274	TRUE
8x8	easy	H2	with_ac3	0.203372	46	TRUE
8x8	easy	H2	without_a	0.013331	90	TRUE
8x8	medium	H2	with_ac3	0.220932	49	TRUE
8x8	medium	H2	without_a	0.014614	96	TRUE
8x8	hard	H2	with_ac3	0.356335	107	TRUE
8x8	hard	H2	without_a	0.029632	212	TRUE
8x8	veryhard	H2	with_ac3	0.31714	81	TRUE
8x8	veryhard	H2	without_a	0.02841	160	TRUE
8x8	easy	H3	with_ac3	6.684766	102	TRUE
8x8	easy	H3	without_a	9.320886	326	TRUE
8x8	medium	H3	with_ac3	21.06807	326	TRUE
8x8	medium	H3	without_a	22.8931	1032	TRUE
8x8	hard	H3	with_ac3	145.7039	6980	TRUE
8x8	hard	H3	without_a	169.6339	23530	TRUE
8x8	veryhard	H3	with_ac3	13.97738	567	TRUE
8x8	veryhard	H3	without_a	13.74677	1274	TRUE

REPORT: ASSIGNMENT 1

COURSE: COSC-4117EL-01: Artificial Intelligence

GROUP: 2

10x10	easy	H1	with_ac3	1.637248	168	TRUE
10x10	easy	H1	without_a	0.03022	896	TRUE
10x10	medium	H1	with_ac3	234.6393	32597	TRUE
10x10	medium	H1	without_a	2.105108	89938	TRUE
10x10	hard	H1	with_ac3			FALSE
10x10	hard	H1	without_a	0.004468	122	TRUE
10x10	veryhard	H1	with_ac3	122.8581	19357	TRUE
10x10	veryhard	H1	without_a	1.520134	72700	TRUE
10x10	easy	H2	with_ac3	0.683245	74	TRUE
10x10	easy	H2	without_a	0.041463	144	TRUE
10x10	medium	H2	with_ac3	0.721487	79	TRUE
10x10	medium	H2	without_a	0.047305	156	TRUE
10x10	hard	H2	with_ac3	0.013106	0	TRUE
10x10	hard	H2	without_a	0.001184	2	TRUE
10x10	veryhard	H2	with_ac3	2.566175	592	TRUE
10x10	veryhard	H2	without_a	0.181226	1176	TRUE
10x10	easy	H3	with_ac3	33.66421	168	TRUE
10x10	easy	H3	without_a	112.122	896	TRUE
10x10	medium	H3	with_ac3			FALSE
10x10	medium	H3	without_a	0.42244	10	TRUE
10x10	hard	H3	with_ac3	0.187221	0	TRUE
10x10	hard	H3	without_a	5.696924	122	TRUE
10x10	veryhard	H3	with_ac3			FALSE
10x10	veryhard	H3	without_a	0.320815	12	TRUE
12x12	easy	H1	with_ac3	14.33479	738	TRUE
12x12	easy	H1	without_a	0.127234	2520	TRUE
12x12	medium	H1	with_ac3			FALSE
12x12	medium	H1	without_a	0.003772	132	TRUE
12x12	hard	H1	with_ac3	68.38911	5146	TRUE
12x12	hard	H1	without_a	0.682713	20078	TRUE
12x12	veryhard	H1	with_ac3	64.7651	5433	TRUE
12x12	veryhard	H1	without_a	1.012888	28160	TRUE
12x12	easy	H2	with_ac3	1.692243	104	TRUE
12x12	easy	H2	without_a	0.074244	206	TRUE
12x12	medium	H2	with_ac3	0.105839	8	TRUE
12x12	medium	H2	without_a	0.004452	18	TRUE
12x12	hard	H2	with_ac3	1.902699	145	TRUE
12x12	hard	H2	without_a	0.095704	286	TRUE
12x12	veryhard	H2	with_ac3	2.960546	291	TRUE
12x12	veryhard	H2	without_a	0.163738	580	TRUE
12x12	easy	H3	with_ac3			FALSE
12x12	easy	H3	without_a	31.39298	86	TRUE
12x12	medium	H3	with_ac3	2.455001	15	TRUE
12x12	medium	H3	without_a	6.090372	132	TRUE
12x12	hard	H3	with_ac3			FALSE
12x12	hard	H3	without_a	1.079543	8	TRUE
12x12	veryhard	H3	with_ac3			FALSE
12x12	veryhard	H3	without_a	2.526636	6	TRUE

REPORT: ASSIGNMENT 1

COURSE: COSC-4117EL-01: Artificial Intelligence

GROUP: 2

14x14	easy	H1	with_ac3	9.885243	347	TRUE
14x14	easy	H1	without_a	0.093167	960	TRUE
14x14	medium	H1	with_ac3			FALSE
14x14	medium	H1	without_a	0.012487	318	TRUE
14x14	hard	H1	with_ac3			FALSE
14x14	hard	H1	without_a	0.000938	10	TRUE
14x14	veryhard	H1	with_ac3			FALSE
14x14	veryhard	H1	without_a	0.000596	4	TRUE
14x14	easy	H2	with_ac3	3.994896	141	TRUE
14x14	easy	H2	without_a	0.171864	278	TRUE
14x14	medium	H2	with_ac3	0.210608	7	TRUE
14x14	medium	H2	without_a	0.006774	16	TRUE
14x14	hard	H2	with_ac3	0.103518	4	TRUE
14x14	hard	H2	without_a	0.004103	10	TRUE
14x14	veryhard	H2	with_ac3	0.054235	1	TRUE
14x14	veryhard	H2	without_a	0.002395	4	TRUE
14x14	easy	H3	with_ac3			FALSE
14x14	easy	H3	without_a	8.09309	10	TRUE
14x14	medium	H3	with_ac3	17.38925	29	TRUE
14x14	medium	H3	without_a	63.72816	318	TRUE
14x14	hard	H3	with_ac3	2.76998	4	TRUE
14x14	hard	H3	without_a	2.984581	10	TRUE
14x14	veryhard	H3	with_ac3	2.08179	1	TRUE
14x14	veryhard	H3	without_a	2.122509	4	TRUE