Benchmark Analysis of TAMER Planner on ICE Problems in ANML

Author Name

June 25, 2025

1 Introduction

This report analyzes the performance of the TAMER planner (Temporal Action Merging Engine for Reasoning) applied to solving ICE (Interleaved Constrained Execution) problems described in ANML (Action Notation Modeling Language). TAMER is a temporal planner designed to handle complex problems with temporal constraints, particularly effective in planning activities that require synchronization and coordination.

ICE problems represent scenarios where multiple activities must be executed with specific temporal constraints and limited resources. The ANML language provides an effective formalism for describing these problems, allowing to specify preconditions, effects, and durations of actions in a clear and expressive manner.

2 Experimental Setup

The benchmarks were conducted using a series of problems of increasing size, identified by the notation "Layer x Item", where:

- "Layer" represents the number of levels or strata of the problem
- "Item" represents the number of elements for each level

For each configuration, different values of the weight parameter (w) were tested, which influences the heuristic of the TAMER planner. The values used range from 0.4 to 0.9, where higher values tend to favor solution optimality at the expense of computational performance.

3 Results

3.1 Solution Makespan

Table 1 shows the makespan (total execution duration) of the solutions found by the TAMER planner for different configurations and values of the parameter w. The theoretical optimal makespan is reported in the second column as a reference.

Table 1: Solution makespan for different configurations and values of w

Layer × Item	**Optimal**	$w = 0.9$	$w = 0.8$	$w = 0.7$	$w = 0.6$	$w = 0.5$	$w = 0.4$
2×3	18t	28	30	28	30	30	26
2×4	24t	38	40	40	40	36	no
2×5	30t	48	50	48	50	46	no
2×6	36t	58	58	60	60	58	no
2×7	42t	70	70	68	70	68	no
2×8	48t	78	78	78	80	no	no
2×9	54t	90	88	no	no	no	no

3.2 Solution Quality

Table 2 presents the execution times of the TAMER planner for finding the solutions reported in the previous table.

Table 2: Execution times for different configurations and values of w

Layer × Item	$w = 0.9$	$w = 0.8$	$w = 0.7$	$w = 0.6$	$w = 0.5$	$w = 0.4$
2×3	305ms	193ms	149ms	130ms	1245ms	37765ms
2×4	1147ms	614ms	463ms	333ms	8564ms	over8min
2×5	3660ms	1778ms	1232ms	911ms	54196ms	over8min
2×6	9004ms	4094ms	2691ms	4281ms	188294ms	over8min
2×7	18435ms	8251ms	12450ms	49417ms	450043ms	over8min
2×8	85948ms	38222ms	23939ms	137148ms	over8min	over8min
2×9	157628ms	67644ms	over8min	over8min	over8min	over8min

4 Results Analysis

From the analysis of the data reported in the previous tables, several considerations emerge:

1. **Trade-off between optimality and execution time**: There is a clear compromise between the quality of solutions (expressed by the optimal/current ratio) and the time required to calculate them. Higher values of w (0.8-0.9) generally produce solutions in longer times compared to intermediate values (0.6-0.7).

2. **Scalability**: As the problem size increases (Layer × Item), execution times grow exponentially, highlighting the intrinsic computational complexity of ICE problems. For problems of size 2×8 and 2×9, many configurations fail to find a solution within a reasonable time (5 minutes).

3. **Optimal value of w**: For small problems (2×3, 2×4), a lower value of w (0.6) offers a good compromise between execution time and solution quality. For larger problems, intermediate values (0.7-0.8) seem to offer the best balance.

4. **Gap from optimal**: In all cases, the quality of the solutions found (expressed as optimal/current ratio) is significantly lower than 1, with values ranging from about 0.6 to 0.7. This suggests that there is still room for improvement in the planning algorithm.

5 Graphical Representation

Figure 1 shows the trend of execution times (on a logarithmic scale) as the problem size varies for different values of the parameter w.

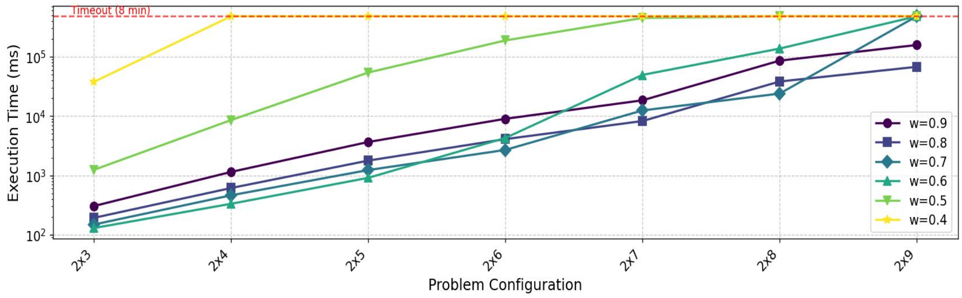

Figure 1: Execution times for different configurations and values of w

Figure 2 shows the quality of solutions expressed as the ratio between the optimal makespan and the actual makespan (closer to 1 is better).

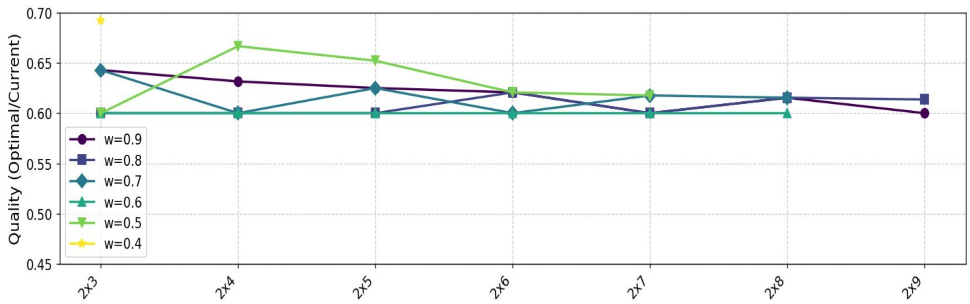

Figure 2: Solution quality (optimal/current ratio) for different configurations and values of w

6 Conclusions

The TAMER planner demonstrates a variable ability to solve ICE problems in ANML, with performance that strongly depends on the problem size and the configuration of the parameter w. In general, it is observed that:

- TAMER is able to find solutions for problems up to 2×9 with appropriate configurations, but with times that grow rapidly as the size increases.
- The quality of solutions, measured as the ratio between optimal makespan and actual makespan, is significantly lower than 1 (the ideal), with values typically between 0.6 and 0.7.
- The optimal value of the parameter w depends on the problem size, with a trend that favors lower values for more complex problems.

Future research directions could focus on improving TAMER's heuristic to bring solution quality closer to the ideal value of 1, implementing more efficient pruning techniques to reduce the search space, and parallelizing the algorithm to tackle problems of larger sizes.