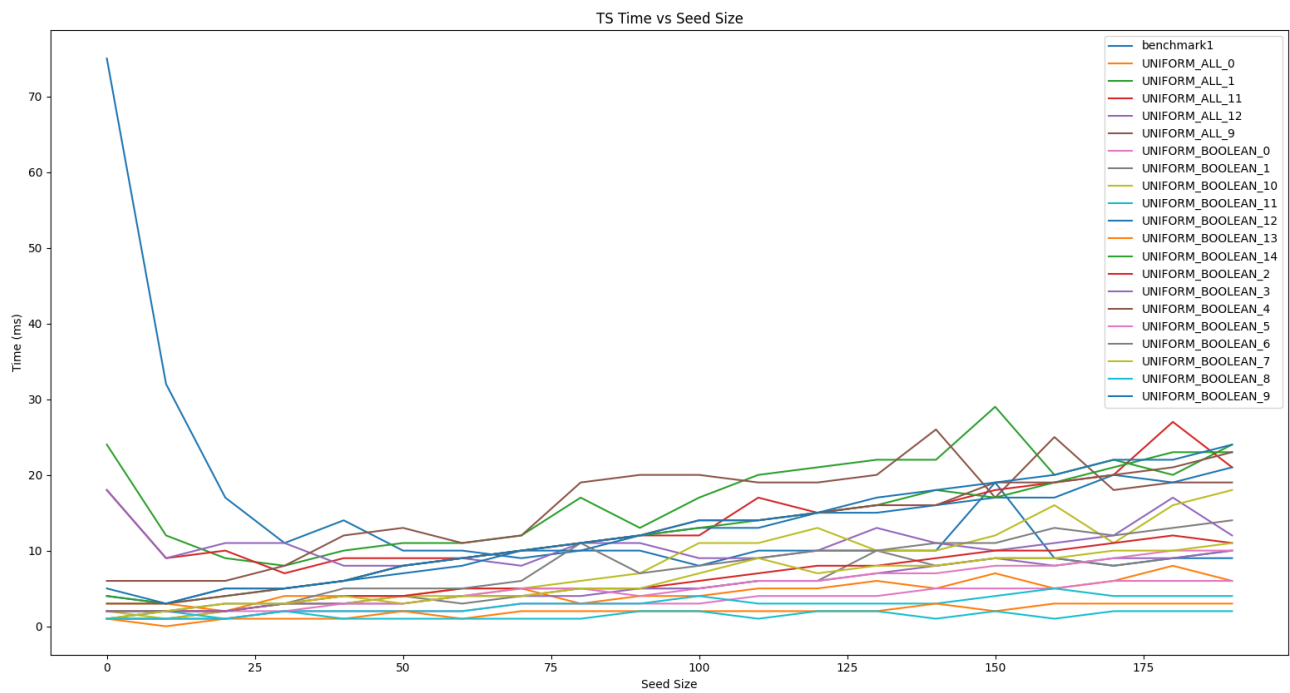
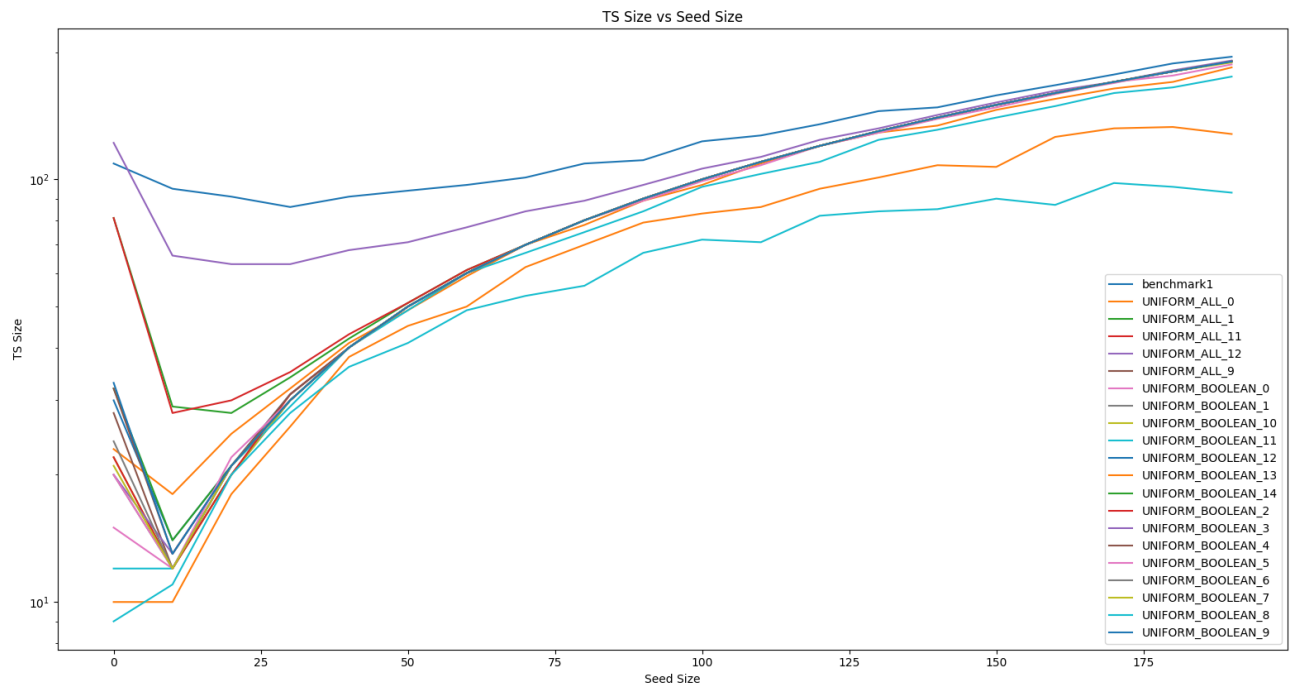


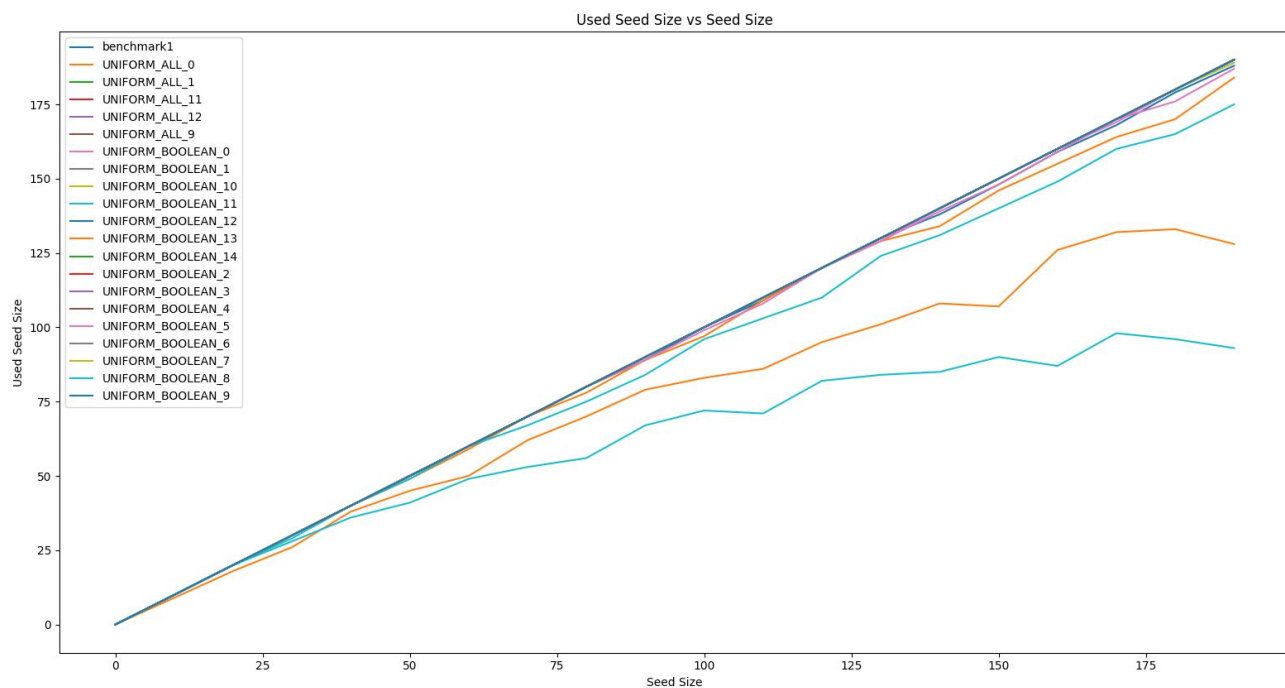
### Mixing RANDOM GENERATION and pMEDICI+

| Model              | Number of parameters | Alphabet size |
|--------------------|----------------------|---------------|
| Benchmark1         | 5                    | 7             |
| UNIFORM_ALL_0      | 7                    | 3             |
| UNIFORM_ALL_1      | 23                   | 3             |
| UNIFORM_ALL_11     | 23                   | 3             |
| UNIFORM_ALL_12     | 12                   | 5             |
| UNIFORM_ALL_9      | 25                   | 2             |
| UNIFORM_BOOLEAN_0  | 13                   | 2             |
| UNIFORM_BOOLEAN_1  | 17                   | 2             |
| UNIFORM_BOOLEAN_2  | 19                   | 2             |
| UNIFORM_BOOLEAN_3  | 17                   | 2             |
| UNIFORM_BOOLEAN_4  | 28                   | 2             |
| UNIFORM_BOOLEAN_5  | 17                   | 2             |
| UNIFORM_BOOLEAN_6  | 21                   | 2             |
| UNIFORM_BOOLEAN_7  | 18                   | 2             |
| UNIFORM_BOOLEAN_8  | 7                    | 2             |
| UNIFORM_BOOLEAN_9  | 29                   | 2             |
| UNIFORM_BOOLEAN_10 | 19                   | 2             |
| UNIFORM_BOOLEAN_11 | 10                   | 2             |
| UNIFORM_BOOLEAN_12 | 27                   | 2             |
| UNIFORM_BOOLEAN_13 | 8                    | 2             |
| UNIFORM_BOOLEAN_14 | 28                   | 2             |

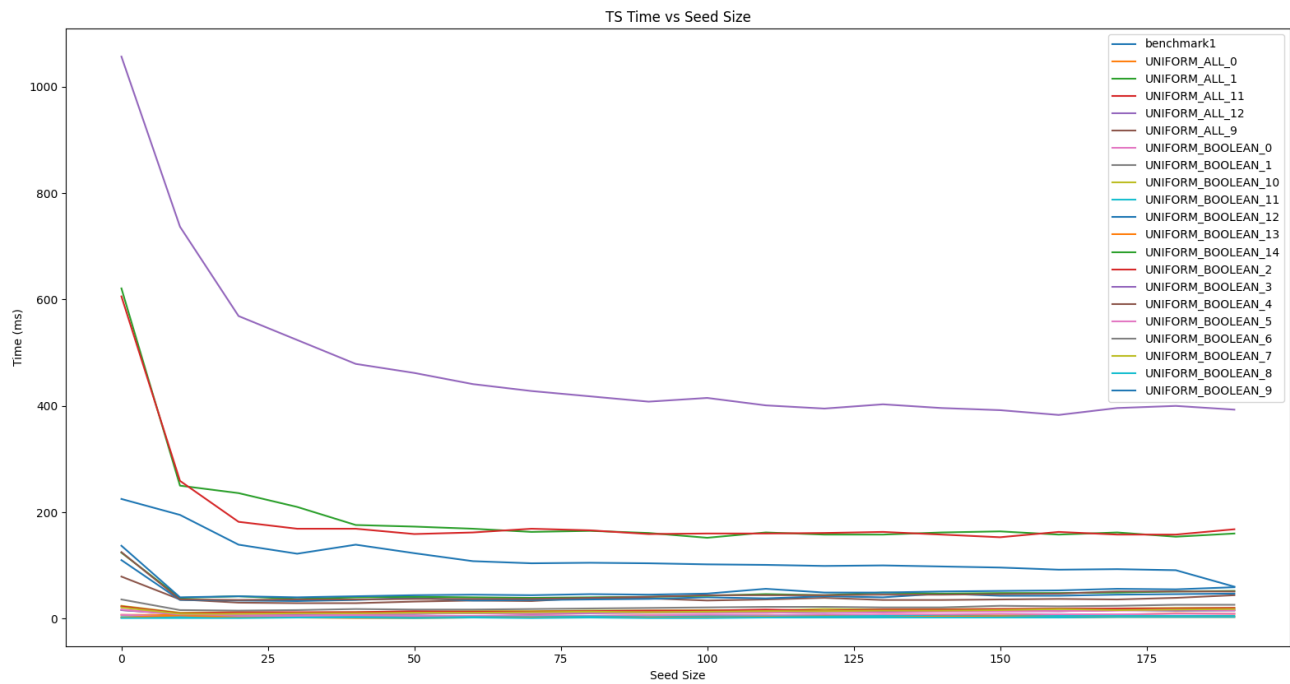
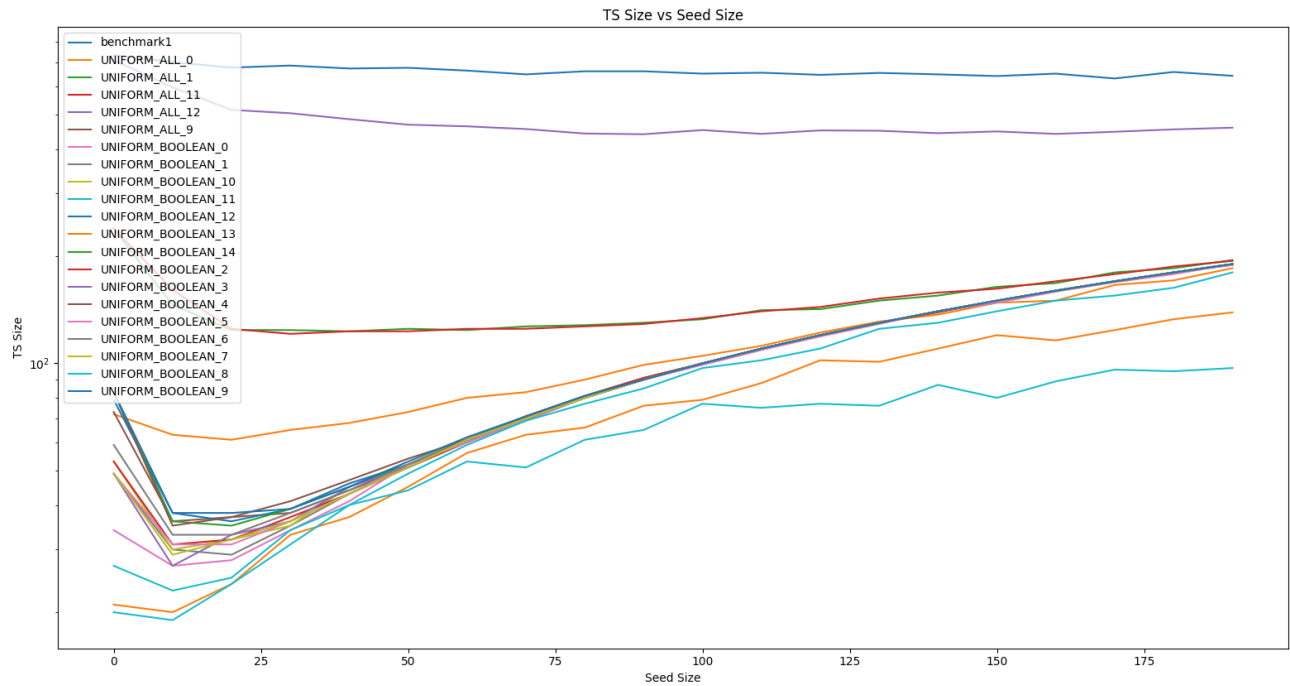
- All models are UNIFORM
- Taken from the last competition (except for Benchmark1) which is taken from the paper “One-Test-at-a-Time Heuristic Search for Interaction Test Suites”
- Seed sizes from 0 to 19 (x10)
- Test suites generated with pMEDICI+ using only one thread in order to evaluate the impact of tests seeding and exclude that of multithreading
- Three measures considered:
  - Test suite size VS Seed size (with logarithmic y-axis)
  - Test suite generation time VS Seed size
  - Actual generated seeds VS Seed size (i.e., when randomly generating CAs, some test may be duplicated and so, the actual generated seeds are fewer than requested)

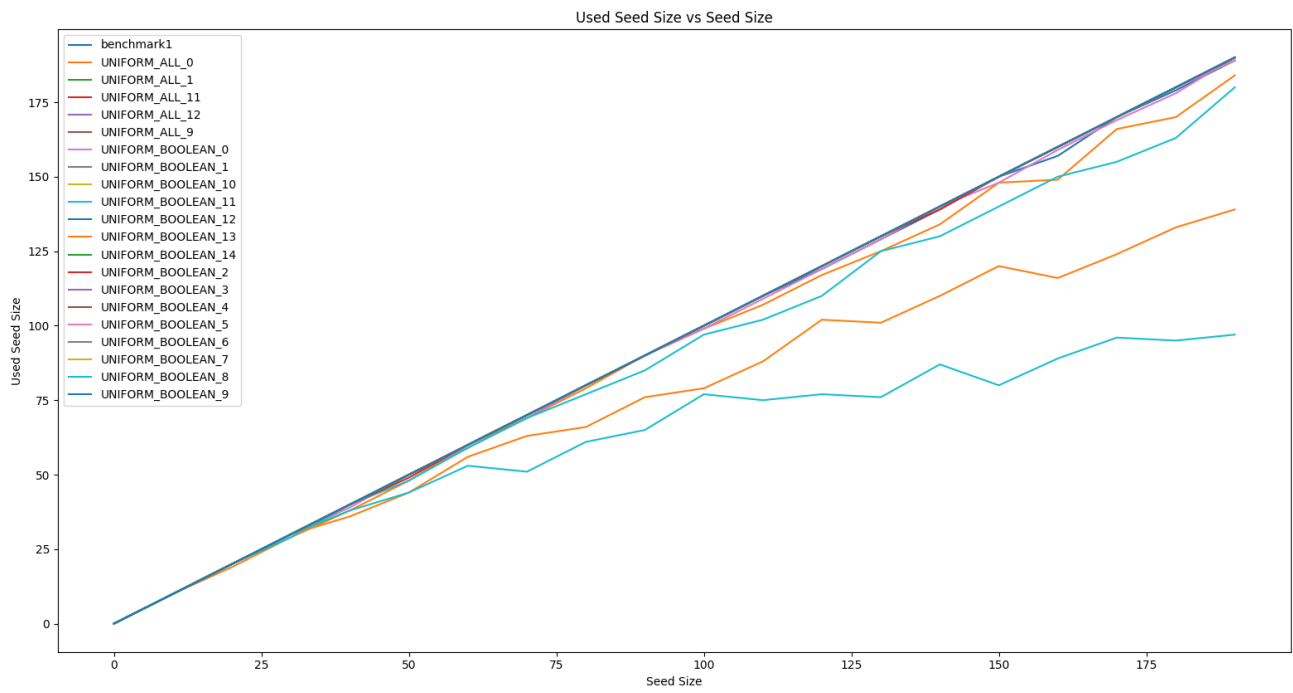
## STRENGTH $t=2$



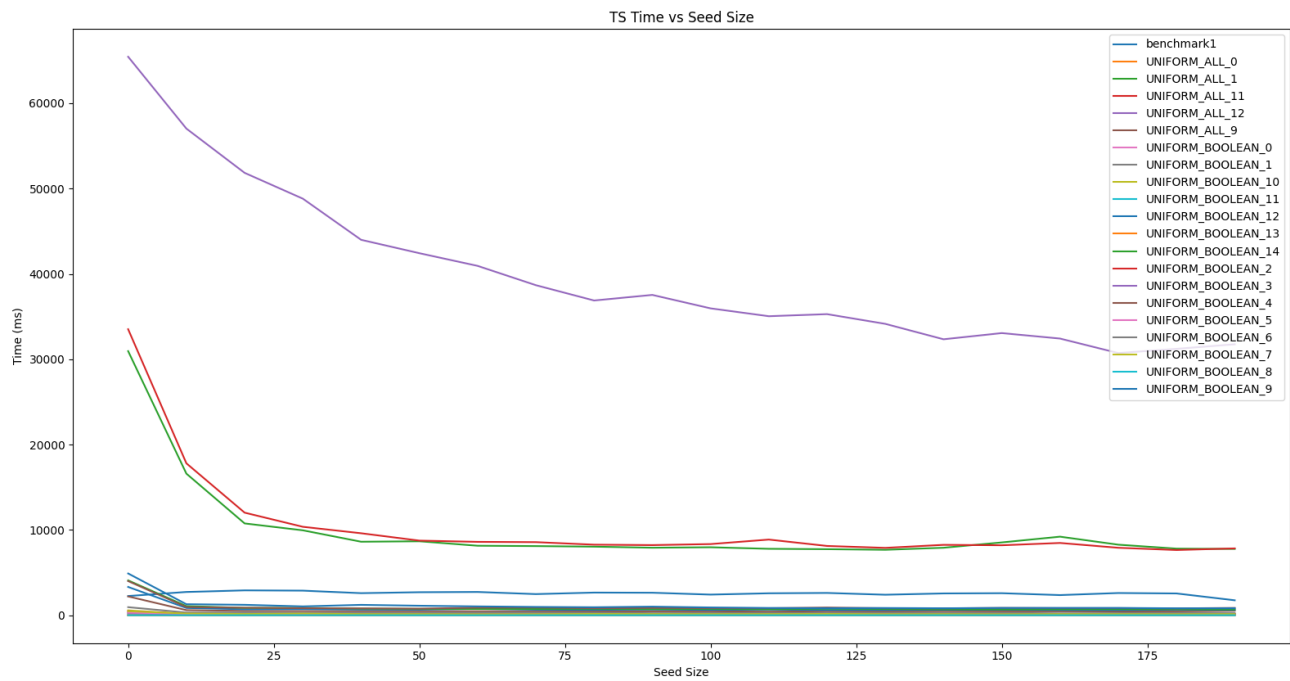
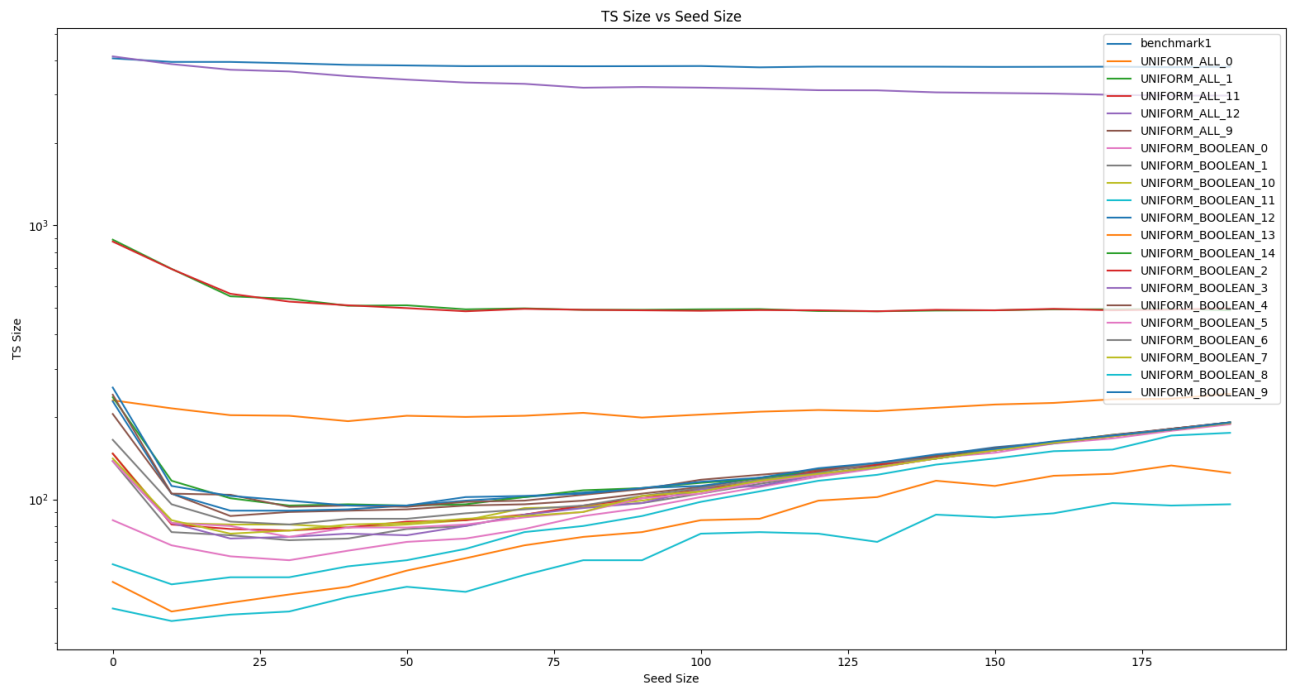


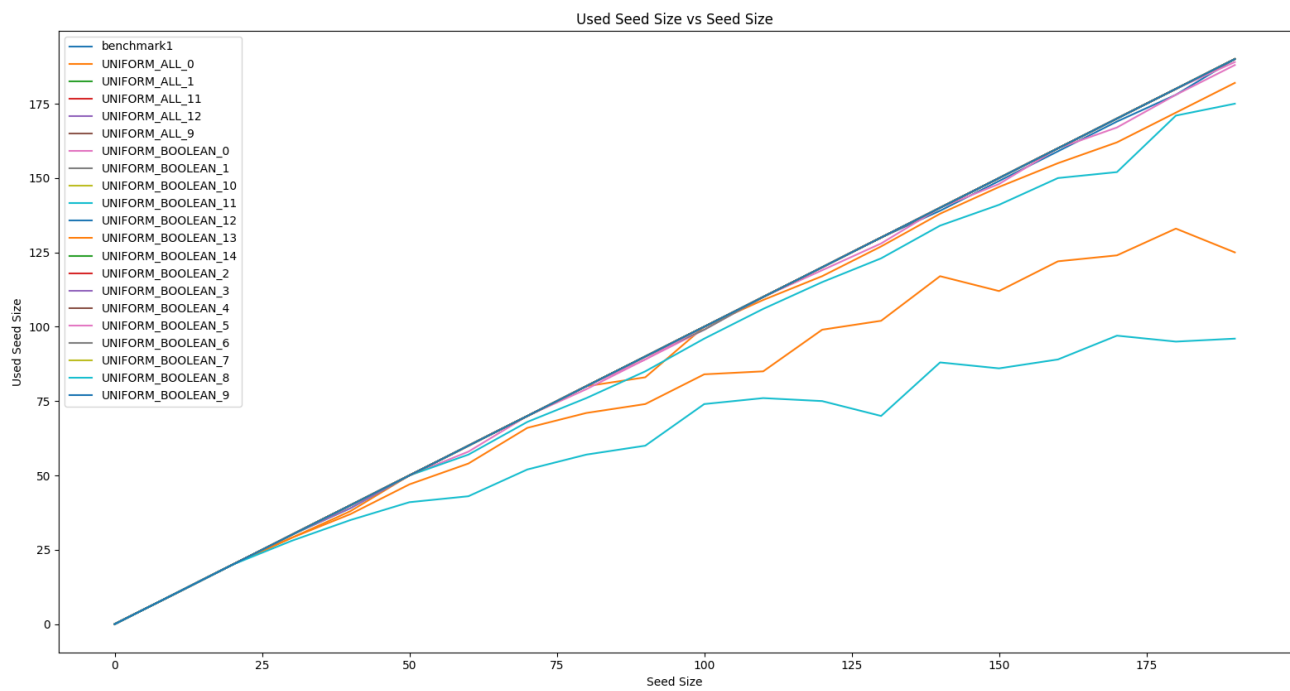
## STRENGTH $t=3$





## STRENGTH $t=4$

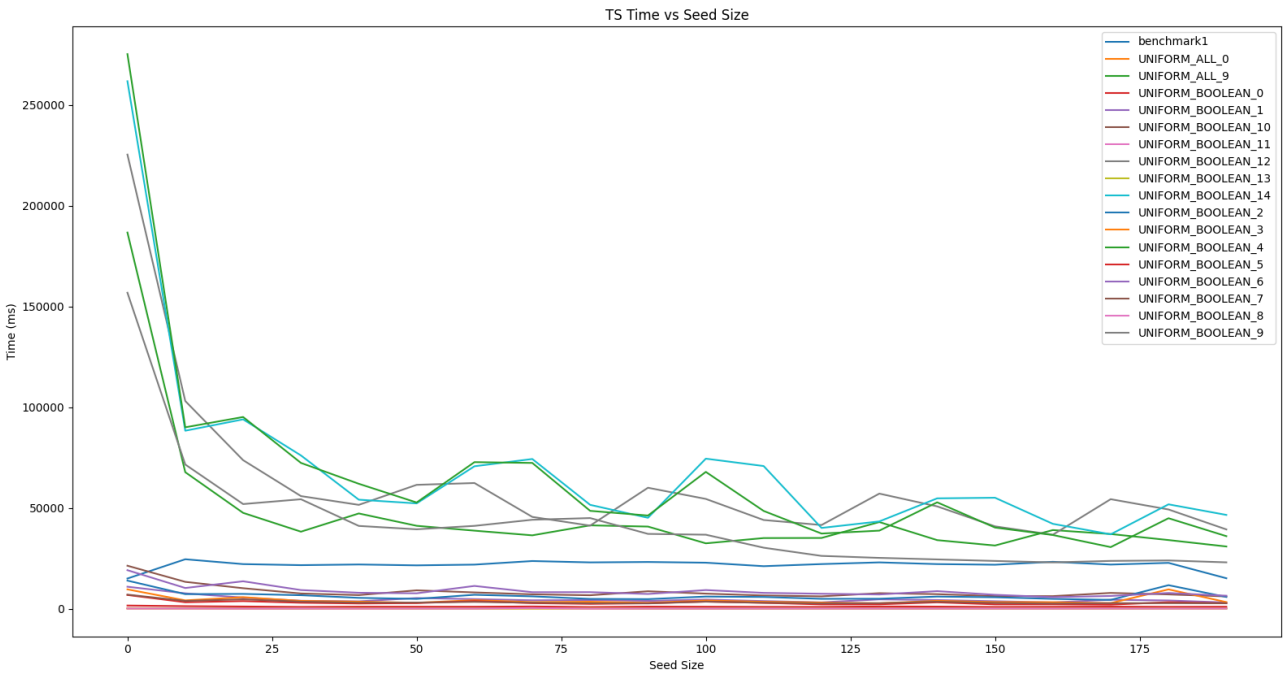
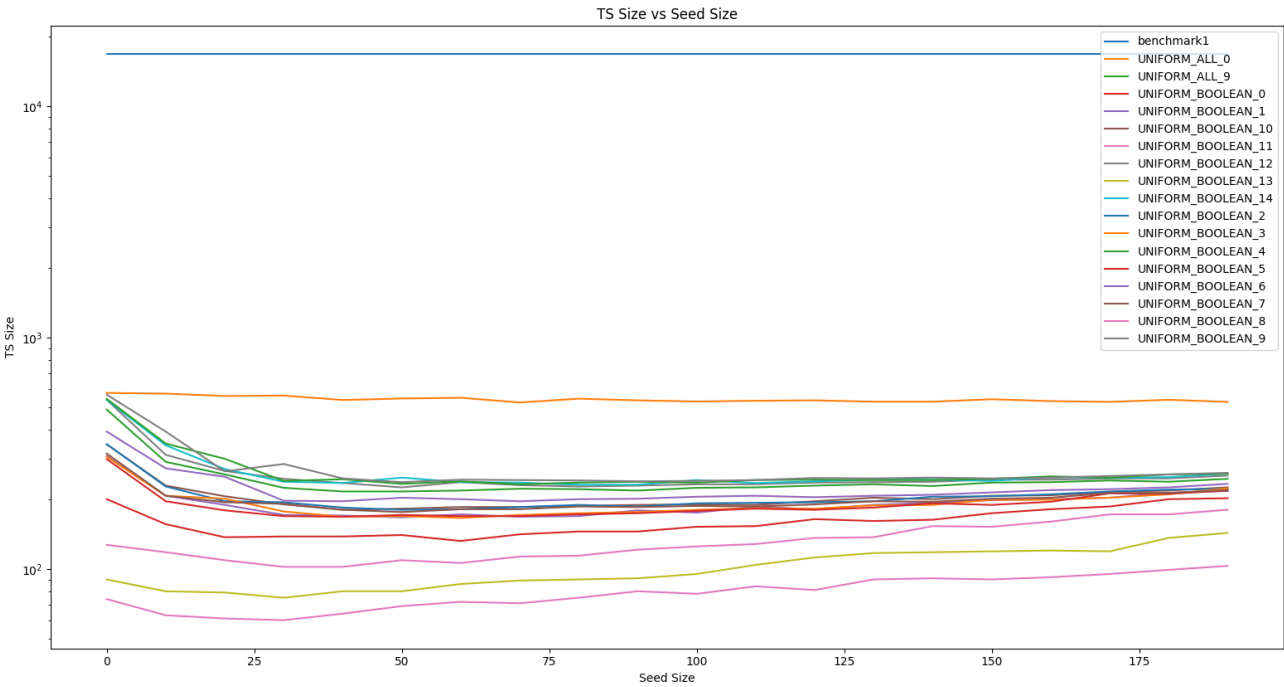


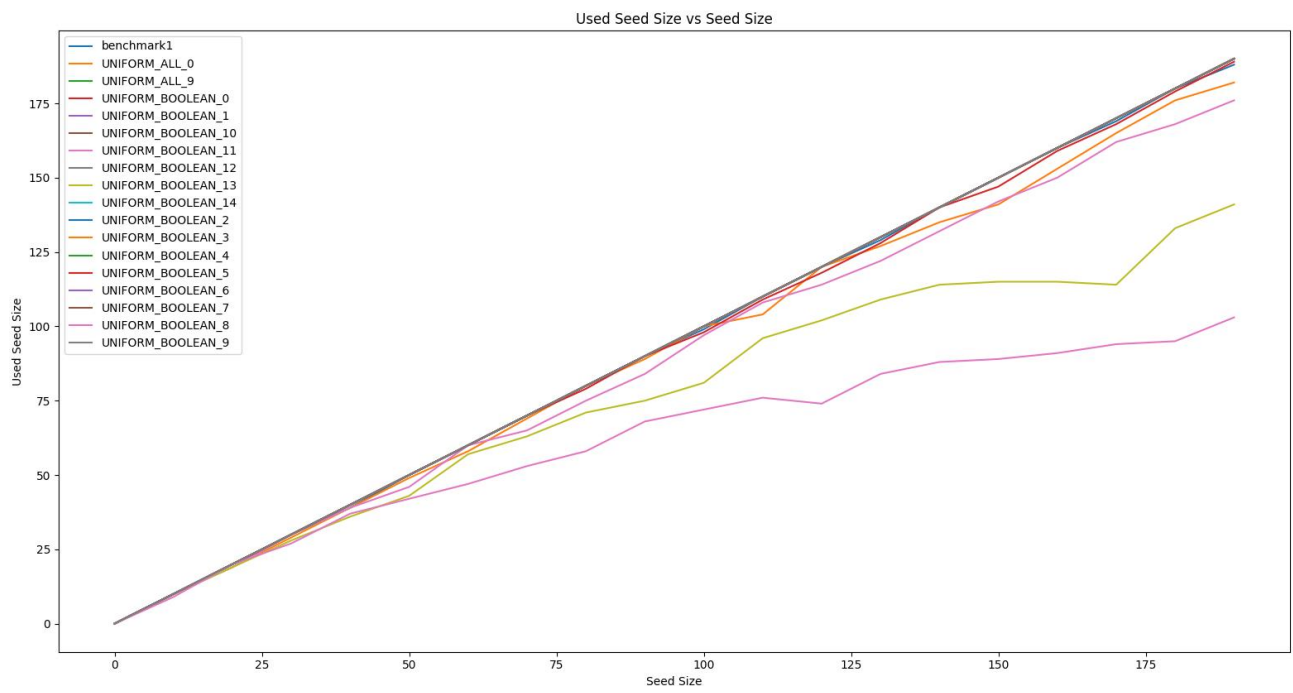






STRENGTH  $t=5$





## **SOME CONSIDERATION**

- In general there is an advantage in terms of generation time when increasing the seed size. The higher the model complexity, the higher the advantage.
- In terms of size, it depends on the number of possible test cases. There is a point beyond which it is useless (I would say counterproductive) to increase the seed size, since there are more random tests than those that would have been generated by the generator starting from scratch.
- For simpler models, which have few combinations, it is possible to have a consistent number of duplicated test cases when randomly generating them.
- It seems that the “best” point for the test suite size is also the “best” point for the generation time.
- Especially for higher strengths the advantage of including a random pre-generation in complex models is very significant especially in terms of time.