

1 A General CA results

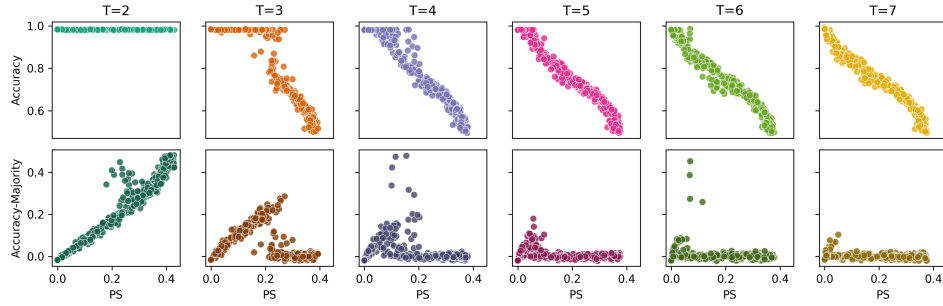


Figure 1: Results of different training runs for different timescales for general automata. From left to right, coarse-graining increases from $T = 2$ to $T = 7$. First row shows accuracy of the trained network, second row shows accuracy difference compared to a simple majority classifier. We can see that with increased temporal coarse-graining it becomes more difficult to learn functions with higher perturbation sensitivity.

2 We perform all experiments in the main text on outer-totalistic automata, as they have additional
3 symmetries, a smaller set of rules and the rules are on average more expressive. Nevertheless, we
4 also provide some results for general automata with a 3×3 neighborhood. One issue here is that
5 random sampling of rules leads to an extreme oversampling of rules with high PS values. We therefore
6 sample rules by first uniformly selecting a lambda parameter (i.e. ratio of patterns that is mapped to
7 value 1), and then sample a random rule roughly fitting that lambda parameter. This provides a much
8 better spread of PS values in our experiments. Using this approach, we again sample 512 different
9 rules, and train networks for $T = 2, 3, \dots, 7$. Results for this temporal coarse graining experiment can
10 be seen in Fig. 1. Compared to the outer-totalistic experiments, we have a much stricter dependency
11 on PS, but at the same time worse overall performance. I.e. in this case we barely have any rules
12 outperforming majority vote baseline starting at $T = 5$, while we still have many such rules in the
13 outer-totalistic case - even up to $T = 7$. The general pattern is the same as in the outer-totalistic case,
14 with accuracy generally dropping with larger T and higher PS.

15 B Additional Plots

16 Here we provide some additional plots for Section 4.3. The given automata were randomly selected
17 using the following scheme: First we make sure that we can even define an object. This is done by
18 checking if all 4 corners of the true label have the same value, as they are the points furthest away
19 from initialization. If yes, this color is selected as the "background" color. We then check if our
20 prediction beats the majority vote baseline by at least 5%. We take the first 16 such automata and
21 plot the results here, showing non-cherry picked results.

22 We see the same pattern as in the main text. In almost all cases, the network is able to predict the
23 outside of the object very well, with a drop as soon as we reach the interior or one or two steps
24 afterwards. In some examples this drop is strong and the network recovers, but in many cases we see
25 a continued drop in accuracy.

