0.1 Determine the Termination Point

0.1.1 Experimental Setup

We choose to run the configuration for evolution cycles from 5 to 100 in steps of 5. We keep the mutation rate at [0.1] for each mutation type. The remaining procedure follows the process described in **??**.

0.1.2 Results

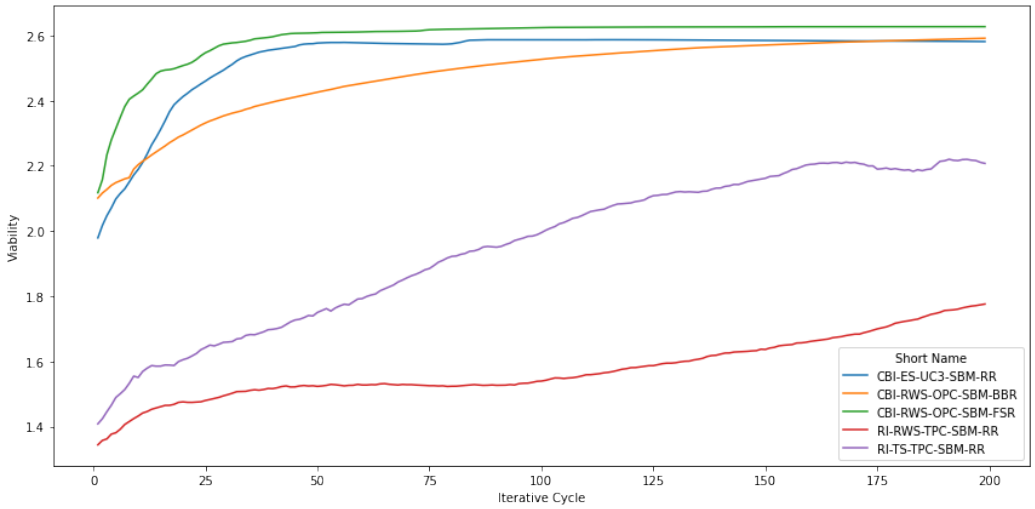

Figure 1: This figure shows the viability of across the iteration cycles.

In Figure 1, we see a general increase in viability for each termination point. It shows, that increasing the termination point also yields better results at the end of the generation process. We see that [CBI-ES-UC3-SBM-RR] returns the best results in the shortest time span. The model converges after roughly 50 iterative cycles. [CBI-RWS-OPC-SBM-BBR] appears to have not reached convergence. The randomly initiated models have not reached convergence as well. However, they remain far below models that use a more sophisticated method to initialize their population.

Figure 2 shows a decomposed view on how the viability measure evolves. Furthermore, we show the average amount of events within a generated counterfactual. In terms of similarity and sparcity all models behave similar. This is no surprise as both measures are ineherently interlinked. We see that the randomly initiated models (RI-x) decrease the amount of events they generate. Case-based initiated models appear to slightly gain more events.

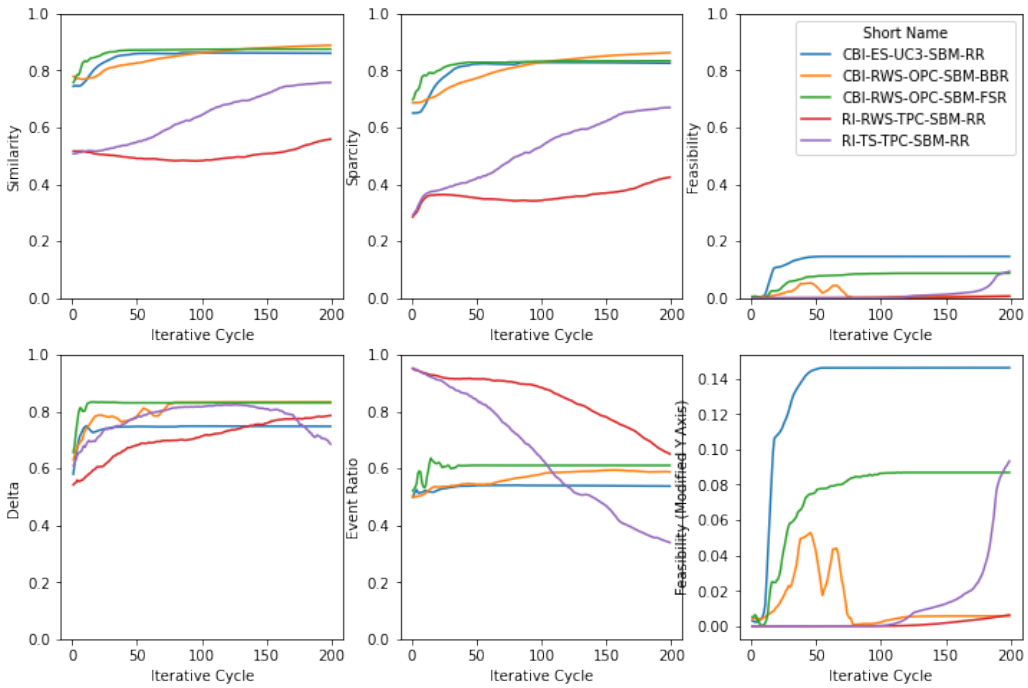

Figure 2: This figure shows the remaining measure components. Additionally, we show the ratio of events within the population. We also show a magnified version of the feasibility measure.

Although, [CBI-RWS-OPC-SBN-BBR] appears that reaches its saturation point significantly later ([100]th cycle). Interestingly, the [CBI-RWS-OPC-SBM-BBR] model struggles to maintain feasibility and collapses to near 0 after the [100th] iterative cycle. Another surprise is the steep ascention of only model that uses tournament selection (RI-TS-TPC-SBM-RR) towards the end of the generation process. The model even overtakes the model that leads the model-configurations in terms of [viability]. Furthermore, we see that CBI-ES-UC-SBM-RR has the highest feasibility among all models. However, it also quickly converges after 50 iterative cycles.

0.1.3 Discussion

The results are not surprising. The longer the algorithm runs the closer it gets to a local minimum. We expect every evolutionary algorithm to converge at some point, as only the best within the population are chosen for the next iteration. If the model does not include enough non-deterministic components the results collapse to one optimal case interms of structure. Hence, the counterfactual activities remain unchanged for the rest of the generation process. The events ratio should optimally approach a number around 0.5 if the factu-

als are evenly distributed in length. All model-configurations seeminly follow this trajectory. However, models ([**RI-TS-TPC-SBM-RR**]) falls below this level. This coincides with its sharp rise in feasibility. We assume this behavior relates to a bias of the feasibility measure towards shorter sequences. The rise and decline of [**CBI-RWS-OPC-SBM-BBR**] shortly before overtaking all other models in terms of similarity and sparcity indicate a trade-off between how close the counterfactual is to the factual and how feasible it is.

For the next experiments we are going to use [**50**] as a termination point. It appears to be a reasonable point in which most models reach their highest viability yield and have not converged yet. We do not seek convergence, as it we want to maintain the diversity of our counterfactuals.