Vic Lankford, Sam Lengyel, Kai Weiner

Replication Report Draft

Methodology of Original Paper

The paper our team has based our project on was about tracking how extinction events impact functional activity for populations of organisms in Avida. Functional activity refers to the ability of organisms to complete predetermined tasks (performing one of nine basic logic operations on one or two random 32-bit strings and then outputting the bitwise correct result). These tasks were the 9 logical tasks: EQU, XOR, NOR, ANDN, OR, ORN, AND, NAND, and NOT. The exact details of this paper can be read at [this link].

Summary of Original Paper's Result

The specific result our team wanted to replicate was the rapid recovery of a population's functional activity after a pulse extinction (extinction happens immediately) that kills off 99.9% of the original population. The program ran for 200,000 updates and the extinction occurred at 100,000 updates. The paper showed functional activity recovering before 10,000 updates after the extinction. The paper did 100 replicates for this treatment. The population structure was toroidal.

Description of Reimplementation and Changes

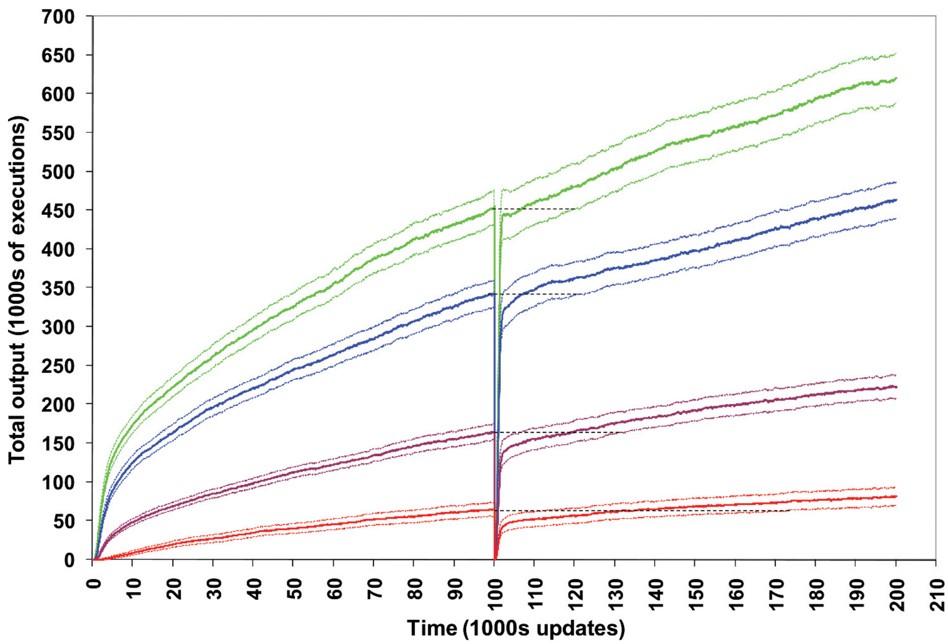
The program from the paper was constructed in Avida. To replicate this project, we constructed a similar project in SGP Lite. We added additional instructions the organisms could perform to imitate those in Avida that were not in SGP Lite - specifically, the IOInstruction and ReproduceInstruction from the Digital Evolution Project sample code and a NAND and NOT instruction (implementation details [here]). We created the tasks used in the experiment - EQU, XOR, NOR, ANDN, OR, ORN, AND, NAND, and NOT (our implementation details [here]). The energy rewards for completing tasks were the same as in the original experiment - 1 unit for NAND and NOT, 2 units for AND and ORN, 4 units for OR and ANDN, 8 units for NOR and XOR, and 16 units for EQU. The original experiment did not specify how many points organisms require to reproduce, so we used 16 units as a baseline amount. We also implemented energy loss every update that would kill organisms who didn't solve tasks to mimic the system in the original experiment. For each CPU cycle, we allowed the organisms one attempt to solve each task. If the organisms had already solved a task in the previous cycle, we did not let them get points from it. If an organism had solved every task in the previous cycle, they would be allowed to try solving each again.

The exact system was not specified so we set organisms to lose 0.125 units for every update they did not solve tasks. The paper also did not specify the threshold at which organisms should be removed. We constructed our program so that organisms which reached -32 units of energy are immediately removed from the population to represent death. During each update, the organisms run 10 CPU cycles, lower than the 30 cycles of the original experiment due to performance issues. We tracked the population of organisms on a 60x60 toroidal grid over 200,000 updates with 99.9% of the population randomly selected to die as an extinction event at update 100,000, the same details as in the paper.

We performed 10 replicates with different random seeds each time - each replicate takes a significant amount of time to run and with the materials available. To perform 100 replicates like the original paper was not possible with time constraints in mind. To begin the world, we seeded it with 8 organisms - more than the original experiment to ensure that the population would be able to fill, with our numbers in mind.

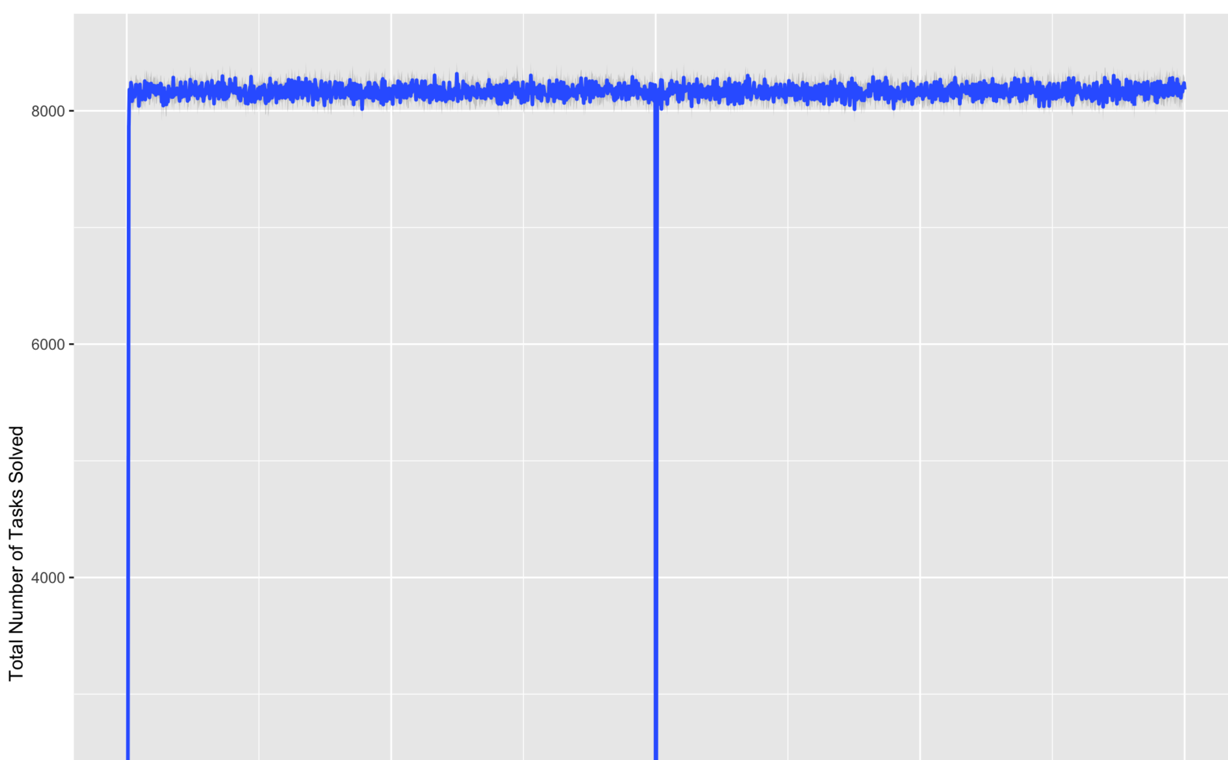
Our Replication Results

Our initial results show a fast recovery of functional activity after the pulse extinction, similar to the paper's results.

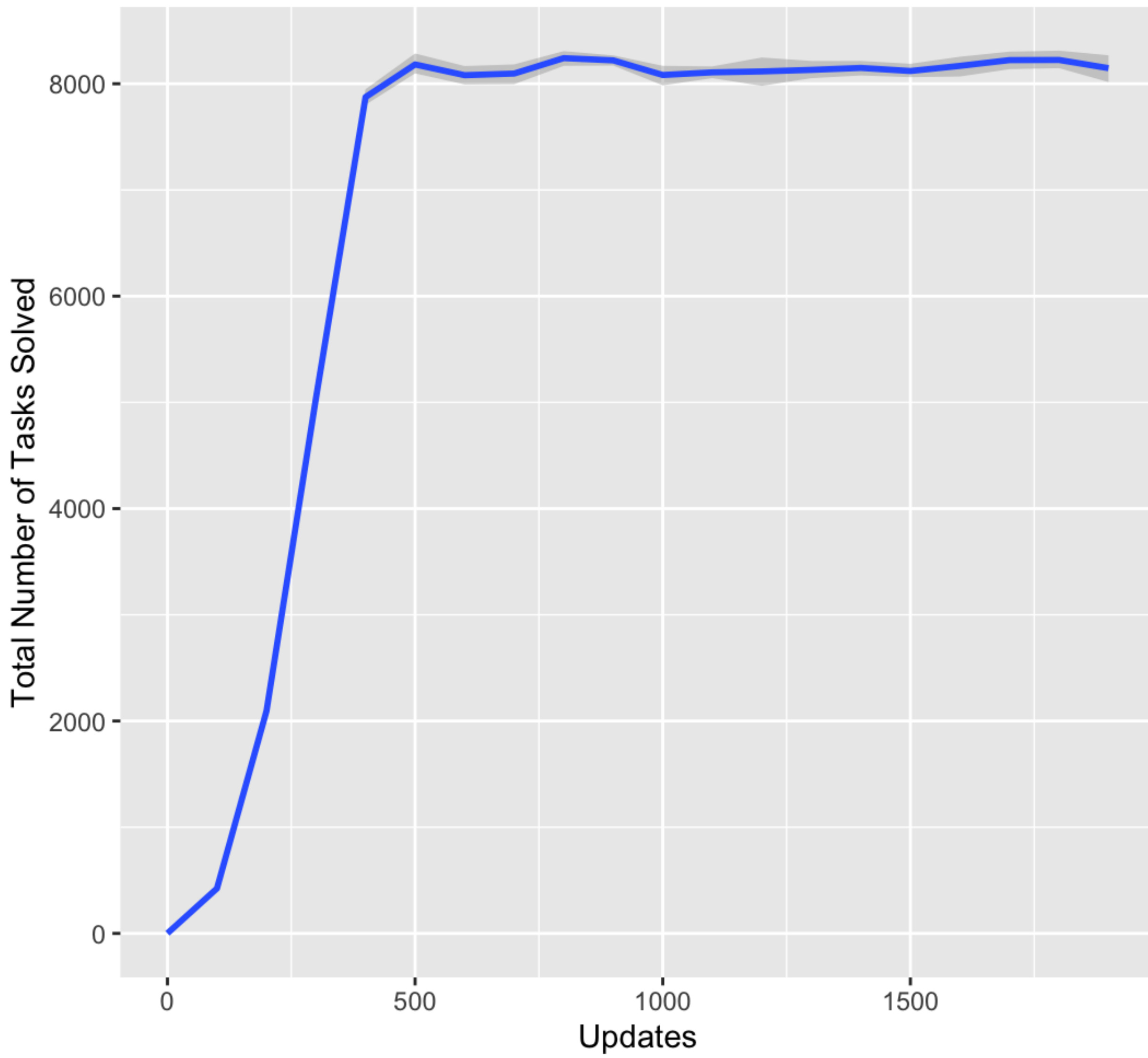


(The paper's results)

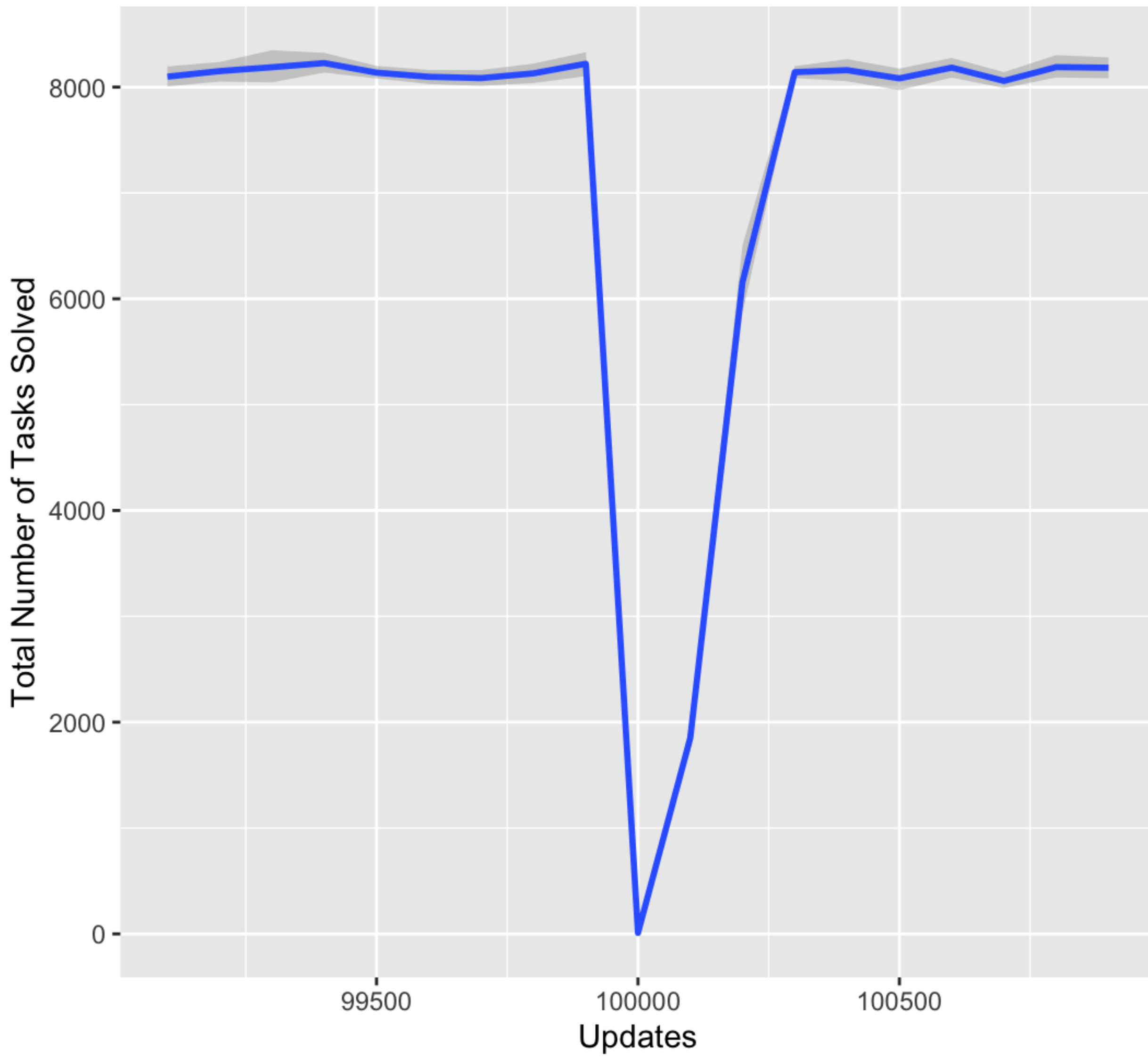
Our organisms should match the red line of the graph because they evolve to perform the highest-level task (Bitwise EQE).



(Our aggregated results over 10 replicates)



Our Results in the first 2000 updates - population increases to having a total number of tasks solved (output) of around 8000, never exceeding 9000 or more tasks solved.



Our results immediately following the extinction - the population recovers to pre-extinction output levels extremely, fast, fully recovering just after 250 updates

Differences in Results

While the overall pattern of functional activity post-extinction is the same in our results as the paper, we observe several differences in specifics. Our program's functional activity increases rapidly in the initial few updates and even faster after the extinction event. After these periods of growth, the functional activity reaches a fairly constant plateau which it oscillates around as time goes on. This plateau is a bit above 8,000 tasks solved by the population per update.

The results of the paper show much slower growth in functional activity. The immediate growth after the extinction event is dramatically faster than the growth at the beginning of the program, even more so than in our program. Functional activity continues to increase with time as the magnitude of growth decreases, no plateau is reached. By the end of the 200,000 updates, the functional activity is still below 100.

We believe that the difference in size of functional activity between our program could be due to the difference in instructions and program used. It is possible that the logical tasks were easier to solve in SGP-Lite or that our method of point rewarding made the programs solve more tasks. This could also be a result of differences in parameters for energy loss and death threshold, our program may have more dramatically selected for organisms that solved more tasks. Intriguingly enough, the difference in CPU cycles would suggest our program would solve less tasks due to less cycles being run meaning less chances for solving. The reduced number of simulations could also have limited the scope of our results, more replications might have brought the functional activity closer to that of the paper's.

The small differences in programs add up. Not knowing the parameter values for energy loss and death threshold means that the selection pressure for more fit organisms between our program and the paper could be dramatically different. This could alter the rate of births, the speed of the population's evolution, and the potential of the population to evolve.

Some of our results seem to suggest that the organisms in our program were not evolving as much as the ones in the paper. The functional activity only changed significantly in periods where the grid was being populated, once population capacity was reached there was not substantial improvement in tasks solved. We observed a greater difference in the growth of functional activity between the initial instances and post-extinction instances of the paper than in our own program. This would suggest there was a greater degree of evolution in the population's ability to solve tasks in the paper's model. A greater ability to evolve would make the post-extinction organisms even better at solving tasks than the initial population. Therefore, the post-extinction population would be faster at accumulating points and be able to reproduce at a faster rate. This means, when given an open grid, a more evolved post extinction population would be able to reproduce faster in comparison to the initial population.

It is also possible that the paper's model would have reached a plateau in growth at some point and that our model's faster growth allowed it to reach that plateau within a shorter time frame. Lastly, some differences between our models could be the result of uncaught bugs in our program, the potential for human error can never be overlooked.

Extensions

The extensions we decided to add to our program were the ability to change the world's dimensions, and allowing the organisms to reproduce pseudo-sexually. The dimensions of the grid where organisms exist can now be altered so that the world can be smaller, larger, and in a rectangular shape rather than a square. In our pseudo-sexual reproduction extension, organisms have the ability to move around and can only reproduce when in the same grid position as another organism. Organisms are assigned one of two sex values denoting their biological identity and are only able to reproduce with organisms of the other sex. In real life, not all organisms that bump into each other end up reproducing, so, to maintain a sense of realism, overlapping organisms have a 50% chance of reproducing. However, the genome of the offspring will be based on only one of the parents with slight mutations, the same style of genetic variation as in the asexual model. There were some conceptual and practical errors in our implementation of pseudo-sexual reproduction which are discussed in our Program Oversights section.

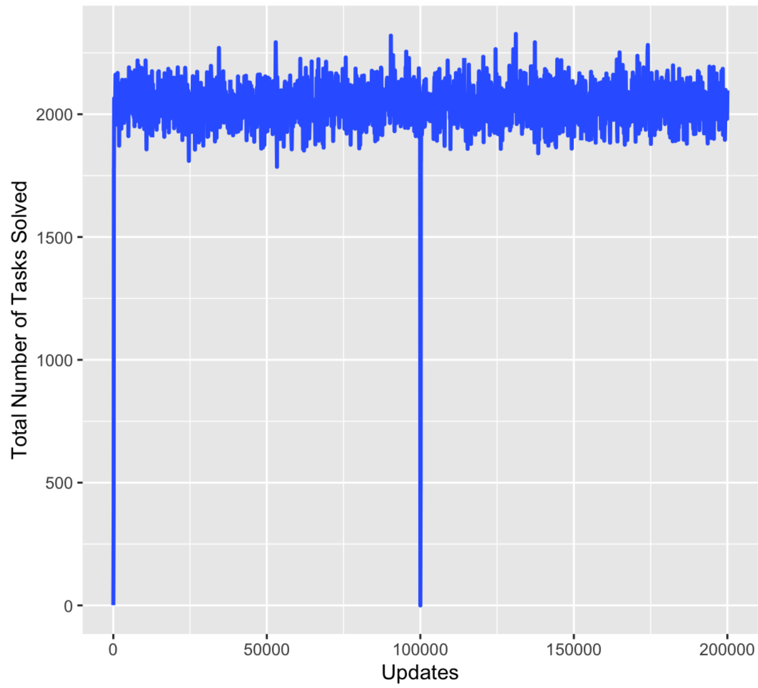
Program Oversights

Collecting data with our program is an intensive process and takes a lot of time. Due to time constraints there are some problems in our code that we caught later in the process and didn't have time to rerun the data on.

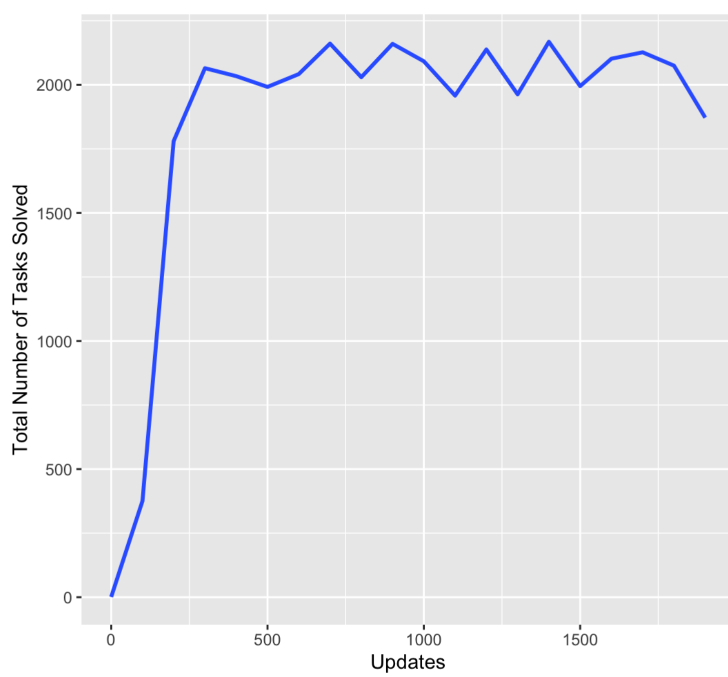
In our movement function for organisms under pseudo-sexual reproduction, we introduced several non-seeded random values to determine gender and chance of two organisms giving each other points. We have since removed these randoms, which could complicate reproduction of data. Additionally, our move function allows organisms to move into the same world position when reproduction is attempted which would cause its attempted mate to be overwritten. We did not intend for pseudo-sexual reproduction to be lethal towards one of the organisms, but may have implemented it as such. Additionally, the reproduction in pseudo-sexual reproduction is actually a 16 point energy boost to the organism attempting to move. This only gives the organism the ability to asexually reproduce later, not actually constituting as reproduction. With these oversights in mind our pseudo-sexual reproduction would be better defined as mutual predation. Organisms or two different types (what we intended to be sexes) are able to kill organisms of the other type and have a 50% chance to gain points from their kill. While not what we intended to model, this is still a useful behavior to simulate and an undeniably funny mistake.

Additionally, there is a discrepancy between the functional activity results in our web and native programs. This does not change our results but does unfortunately make them slightly harder to visualize in our web program.

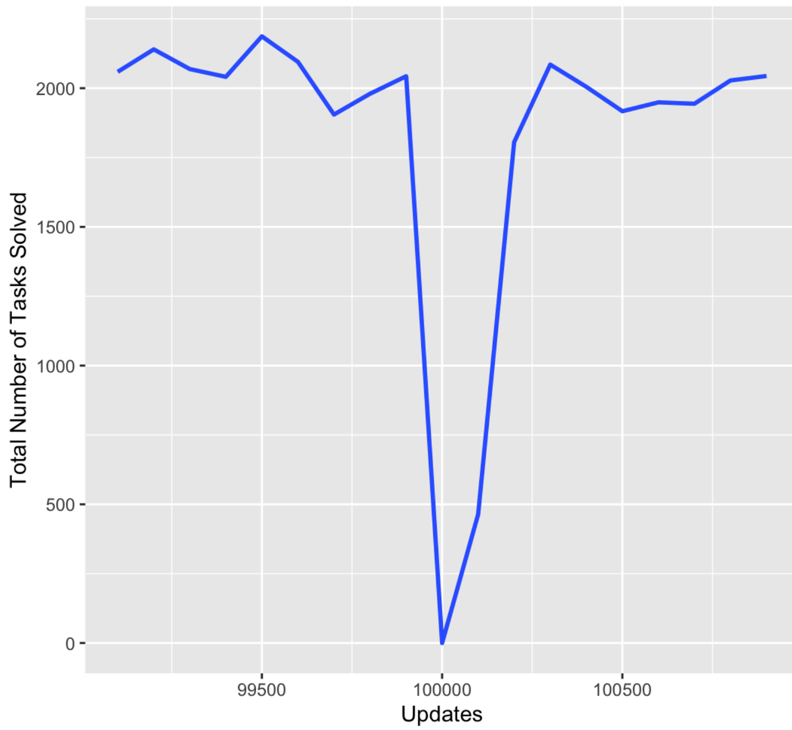
Results with Changing World Size Extension (world size changed to 30x30):



(Our aggregated results over 10 replicates)



Our Results in the first 2000 updates - population increases to having a total number of tasks solved (output) of around 2000, never exceeding 3000 or more tasks solved, and dips at around update 1800.

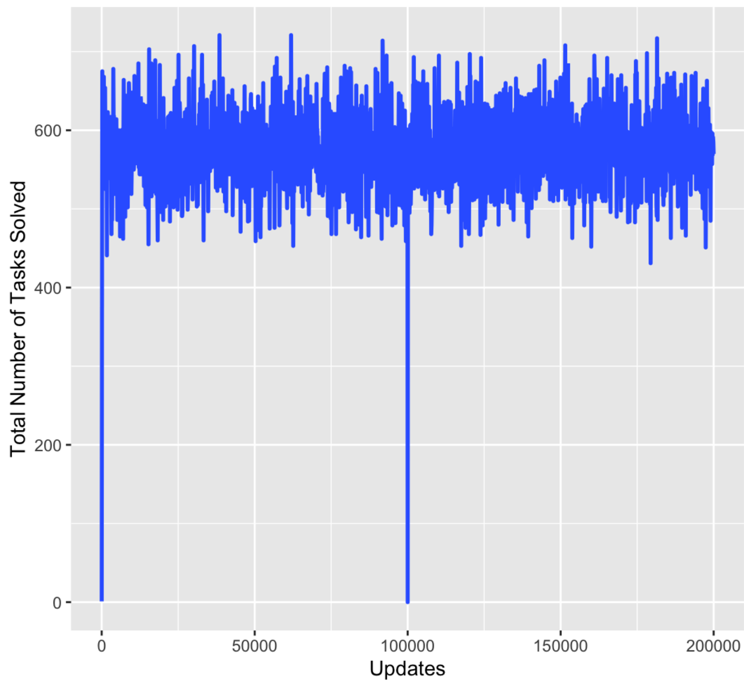


Our results immediately following the extinction - the population recovers to pre-extinction output levels extremely, fast, fully recovering just after 250 updates

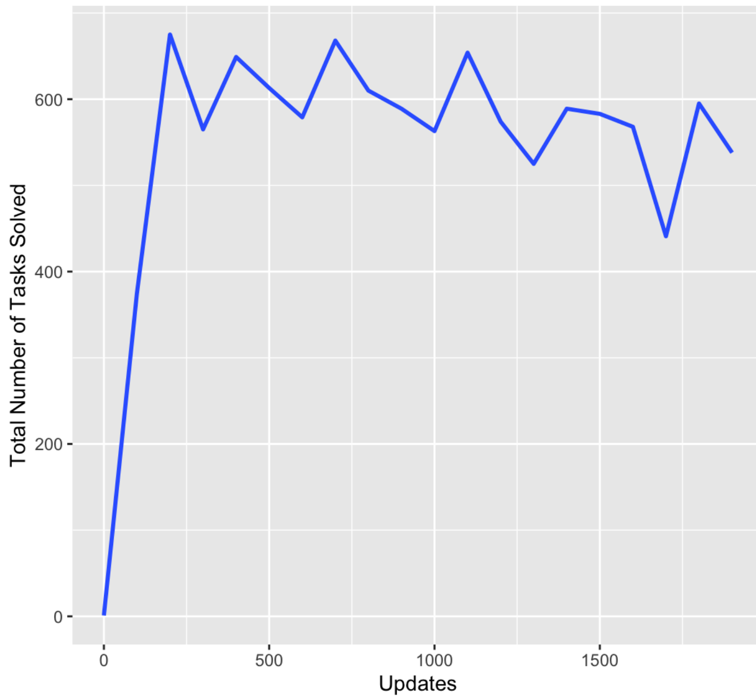
Running the simulation with a smaller world size essentially produces a scaled down version of the original simulation. Initial and post-extinction growth is still rapid, now both are faster and the initial growth period is about identical to the post-extinction growth period. The smaller world size could be the cause of this difference, allowing organisms to reach maximum capacity at a faster rate and thereby reach the functional activity plateau faster. To clarify, this is because a greater number of organisms are able to complete a greater number of tasks. The smaller maximum population may also limit the genetic diversity and amount of evolution the organisms undergo, possibly explaining the smaller difference between growth periods. Additionally, the lower value of the plateau means it can be reached faster.

A plateau still occurs but it is at a significantly smaller value than the initial simulation. This is likely due to the smaller population capacity. A 30*30 grid allows for 900 possible organisms which is 4 times less than the 3,600 possible organisms in the initial simulation. Supporting this idea, the functional activity of the 30x30 world simulation is the same ratio of about 4 times smaller than the initial simulation (plateau around 2,000 for 30x30 world, plateau around 8,000 for the 60x60 world).

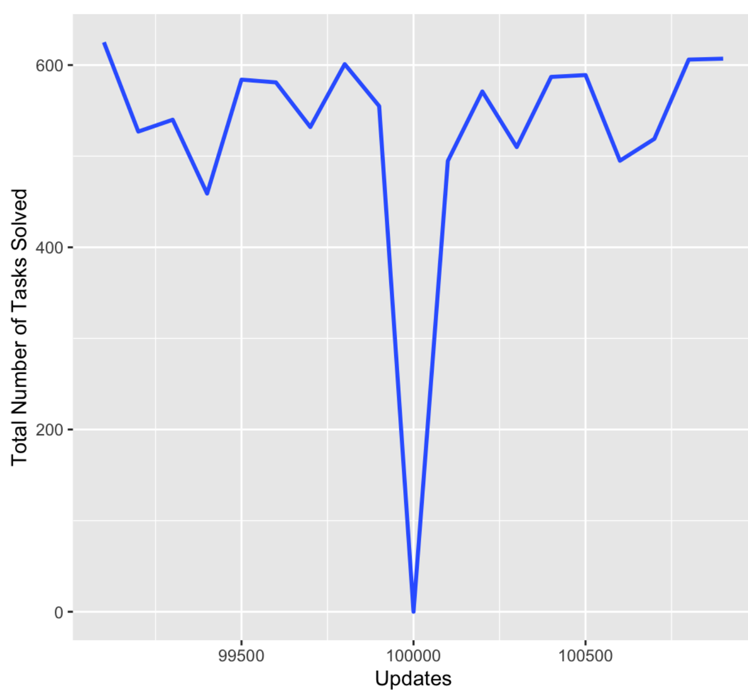
Results with Pseudo-Sexual Reproduction Extension (pseudo-sexual reproduction enabled):



(Our aggregated results over 10 replicates)



Our Results in the first 2000 updates - population increases to having a total number of tasks solved (output) of between 700 - 500 tasks solved.

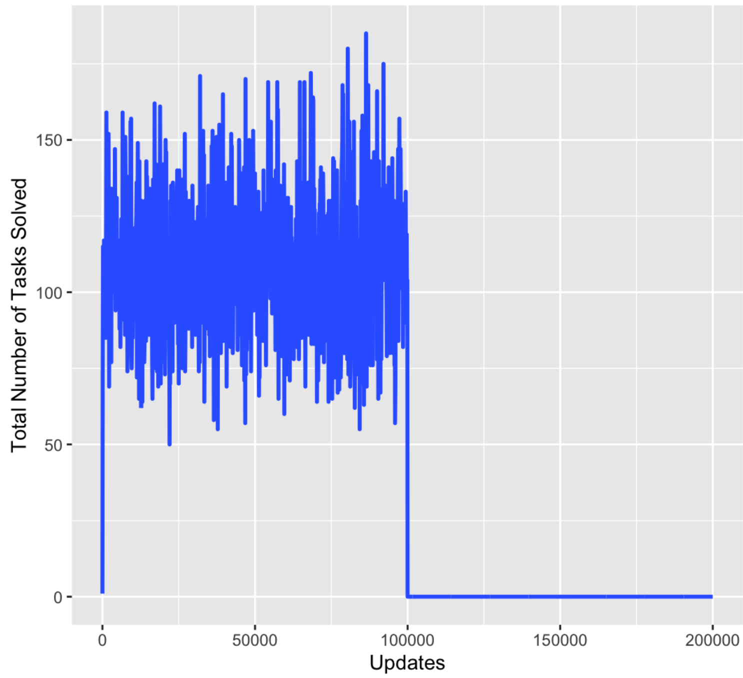


Our results following and leading up to the extinction - the population recovers to pre-extinction output levels fully recovering just after around 250 updates

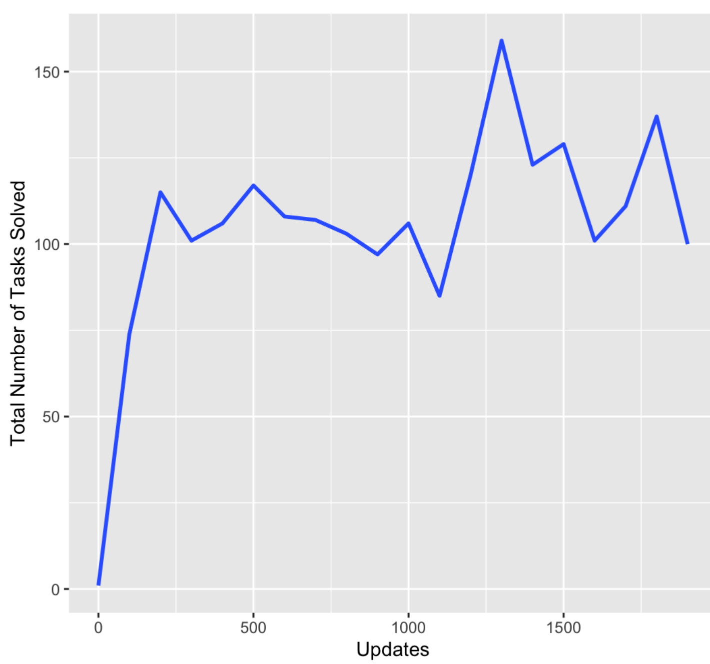
Running our simulation with pseudo-sexual reproduction has several effects on our results. We observe a greater degree of oscillation around the plateau, the plateau value is lower than that of the original experiment, and growth at the beginning of the simulation and immediately after the extinction reaches the plateau much faster (the post-extinction growth period appears faster as in the original).

As described in our Program Oversights section, organisms under pseudo-sexual reproduction kill their attempted mates. As a result, the maximum population size will not be reached due to organisms limiting themselves which could explain the lower functional activity we observe due to not having as many organisms to complete tasks. Population size may vary during the simulation, which would explain the greater degree of oscillation in functional activity. Similarly to the smaller world simulation, the functional activity plateau is likely reached in a faster timespan because it is at a lower value.

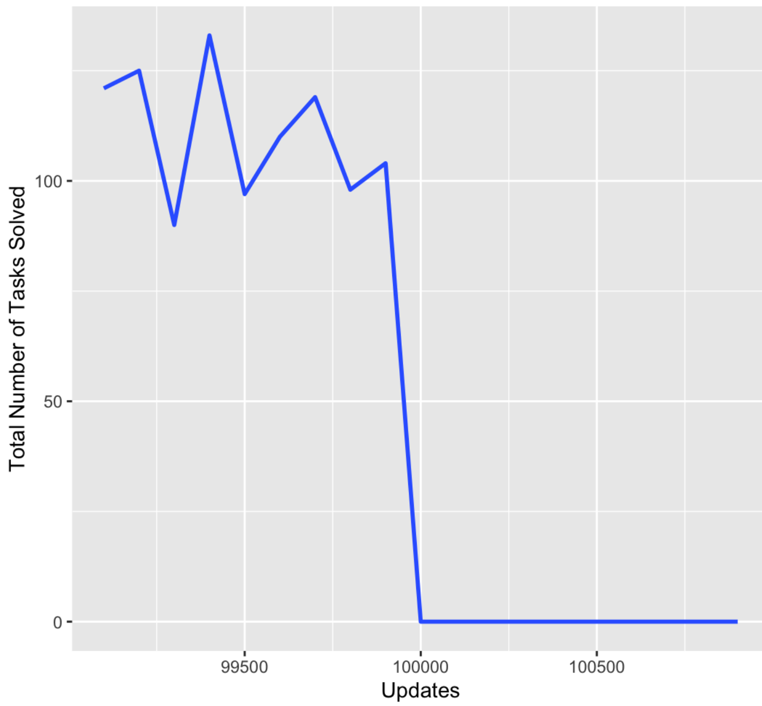
Results with Both Extensions (world size changed to 30x30, pseudo-sexual reproduction enabled):



(Our aggregated results over 10 replicates)



(Our Results in the first 2000 updates - population increases to having a total number of tasks solved (output) of around 150. The number of tasks solved seem to stay between the range 90-110 up until 1000 updates exponentially increases to around 160 tasks solved.)



(Our results immediately following the extinction - this time the number of tasks solved does not recover to its original mean value. This is mostly due to the fact that organisms do not recover past 100000 updates.)

Combining our extensions leads to a dramatically different result to the original simulation. The initial growth period is much faster, the plateau is at a much lower value, and the functional activity does not recover after the extinction event. The lower functional activity makes sense from our previous two results, having a smaller world and implementing pseudo-sexual reproduction both separately decrement the functional activity reached so their combination should also decrement functional activity. In the same vein, a lower plateau would be faster to reach as seen.

The failure to recover functional activity after the extinction could mean that the population was not able to reproduce back and went extinct. The smaller world size in combination with pseudo-sexual reproduction means the population at extinction time was likely much smaller than in the original simulation meaning the 99.9% extinction event likely left very few survivors. These survivors were likely evolved to depend on interactions with other organisms for energy points and may not have been as effective at solving tasks. As a result, they may not have been able to find other survivors (if there were any) in the smaller grid and died of starvation. A population of zero organisms would have zero functional activity.

Future Extensions

Some small future extensions we might pursue for this project if we had more time are: rerunning our data collection with our updated code, running the data on more cycles to more closely match the original program, and trying other combinations of parameters to more closely mirror the paper's results.

It would also be interesting to perform our simulations on a more biologically accurate version of sexual reproduction that reconbines genomes and doesn't kill potential mates (unless we were simulating praying mantises or animals that follow such behaviors). We could also lean into the mutual predation we may have accidentally created and explore food chain and ecosystem dynamics in our extinction simulation.