

Project Abstract and Timetable – NLP
Zakhar Manikhas (342663978) and Sergiy Horef (341332336)

Abstract

Recent progress in development of Large Language Models (LLMs) has motivated researchers to create LLM-based agents for interactive tasks involving human and artificial agents.

As LLMs have shown to be efficient at predicting and simulating human-created text, later studies suggested techniques for evaluating decision-making strategies, concentrating on language-based persuasion games, in which the agent aims to influence others' choices through verbal communication.

Following research (Eilam, Reut, Moshe, Roi 2024) showed how using simulation could improve performance of models trained to predict human's behavior.

This paper is mainly interested in making changes upon the simulation scheme implemented in the aforementioned paper by combining the strategies proposed there with the work done by Robert Axelrod (1980). Even though Axelrod has mainly worked on the Prisoner's Dilemma, he has proposed a set of rules which empirically produce strategies that are robust in many game-playing scenarios.

Moreover, some of the winning strategies described in his 1980 paper may be finetuned to fit the sender-receiver game setting.

We try to show how the use of more complex strategies (like TIT for TAT), may produce a behavior closer to that of a human player (in comparison to the basic strategies shown by Eilam, Reut, Moshe and Roi), and therefore increase the performance of the simulation augmented model.

Workplan

We have to give in the work on 10.6.

We have very important tests on dates 11.4 and 12.4, therefore we will not progress in our work until then.

In total it leaves us with around 60 days of work, not counting the days in which we do other tests.

We would like to spend 1 to 2 days researching the subject of additional strategies as we have explained earlier in the abstract. And afterwards we would like to try 3 to 4 different strategies – two of them being TIT for TAT and Majority rule. One or two others will be fitted from the strategies presented by Axelrod (1980).

We will therefore have around 58 days to run the code.

We would also like to spend 1 to 2 days writing the final paper and making conclusions from the results we have got. And include 2 days spare in case something goes wrong in any part of our work.

This leaves us in total with around 54 days of work.

Single run of the sweep (with a single basic nature option) provided to us takes around 3600 minutes to test all the hyper-parameters provided. That is, around 60 hours, or 2.5 days. This number is only estimated from the fact that each single run of the StrategyTransfer file with default parameters takes around 2 hours to finish, because we were not able yet to run the sweep code until the end.

We will still refer to 30 hyper-parameter optimization runs as sweeps.

We have tested running StrategyTransfer file in parallel – running 6 different variations results only in 1 hour of an overhead. That is, if we run 6 different configurations together, we finish all in under 3 hours, instead of the 2 hours for a single run. Hypothetically it would be possible to run 10 configurations until the CPU on the VM becomes a bottleneck, but even 6 is more than enough.

In this way (running 6 in parallel) it would take us around 90 hours, or 3.75 days for 6 sweeps.

If we continuously run the code for all the 54 days we have, we can run approximately 87 sweeps, or, in other words – check 87 different options for basic nature, in the best case.

As we would like to try and test each strategy by itself, and every possible combination of the original 2 strategies (not counting the oracle; that is – truthful and random) and our additional 4 strategies (that is – 6 strategies altogether), we have (in the upper limit): $4+64=68$ sweeps needed.

If we subtract the combinations already checked by the original paper – around 4 combinations in total, we are left with 64 sweeps to check.

This number is much less than the total number of sweeps we can possibly do, so we will probably be able to finish running everything on time.

The final timetable looks something like this:

9.04 – 14.04: researching the strategies and setting the sweeps to run.

14.04 – 4.06: running as much sweeps as we can.

4.06 - 10.06: finishing everything up and writing the final paper.