Self-Generated Culture

François Fleuret

July 8, 2024

(work in progress, to be updated)

https://fleuret.org/public/culture/culture.pdf

1 Introduction

The hypothesis behind this experiment is that high-level abstract thinking is fueled by social competition. A group of communicating agents that try to demonstrate their cognitive superiority would end up developing a rich and consistent culture.

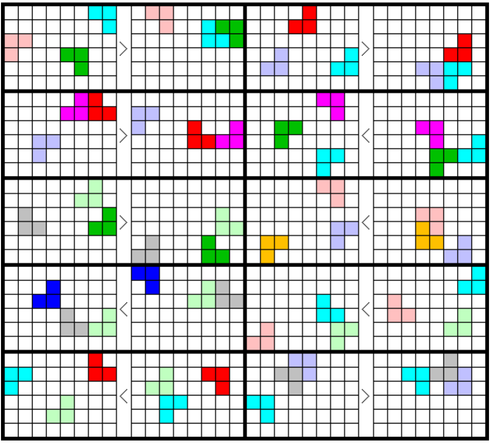
The experiment is designed with a group of GPTs that alternatively learn to solve quizzes and generate new ones.

A "quiz" is a triplet of the form (A, d, B) where A and B are two sequences and d is a token indicating if the direction is forward or backward. Given (A, d), the challenge is to generate B.

The experiments starts with a set of quizzes, that is going to be progressively enriched.

2 Bird World

The initial set of quizzes consist of predicting the dynamics of a very simple world: A 6×8 grid with three colored "birds" moving in a straight line, possibly bouncing on the grid's borders. There are ten different colors.



In each on these quizzes, A is the left image serialized in raster-scan order as a sequence of $6 \times 8 = 48$ tokens, d is either the token "forward" or the token "backward", and B is the right image, also serialized. The direction of prediction is chosen at random.

3 Generating Quizzes

Given a set of N GPTs, we can generate new quizzes as follows: Select one of the models, and use it to generate the 97 tokens of a triplet (A, d, B).

Then with each one of the $N-1$ other models, predict B from (A, d), and A from (B, d') where d' is the direction token opposite of d.

A quiz is validated if **all the other GPTs but one predict it deterministically correctly in both directions.**

This criterion assures that the new quizzes are both solvable and sophisticated, and incrementally complexify the culture. Imposing both direction prevents the generation of quizzes which are not trivial only because the prompt has been randomly degraded.

4 Overall Process

The overall process consists of training the GPTs from scratch by iterating the following steps:

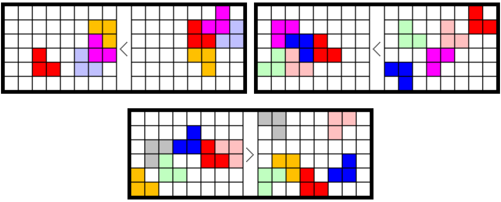
- select the GPT with the lowest recorded test accuracy, train it through one epoch,

- if its test accuracy gets above 97.5%, generate 1′000 new quizzes, add them to the training set, re-compute the accuracy of all the models

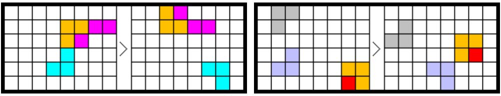
5 Results

This procedure results in the discovery of patterns which are not present in the original quizzes:

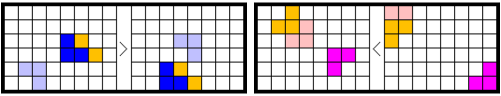
More birds



New bird shapes



Occlusions



6 Various thoughts

- The whole process can be envisioned as natural selection of quizzes in the representation landscape of GPTs. There probably is a subtle relation between the temperature (mutation rate) and the number of models used to validate with the "all but one" criterion (survival criterion).

- The "all but one" could be "all but K", and there may be some information-theoretical thing, where the goal is to maximize mutual information, with $K = N$ being total randomness, so high entropy but no structure, and $K = 0$ is total determinism, so no information to share.

- The setup does not push toward any specific invariance or property in the generated quizzes, their consistency is entirely due to the statistics of the "world quizzes" that remain in the training set, and to the GPTs' inductive biased.

- The GPTs obviously get a sense of objectness and 2d topology early on, since they rapidly increase the number of birds and "discover" occlusion even though they never was in the world quizzes.

- There may not be so many problems that can be cast as pairs of patterns that are each a deterministic function of the other, which is probably critical here.

- This overall process probably fight the "simplicity bias": If a model is lacking a "cue" that the others have, there will rapidly be quizzes that require this cue, they will be added to the training data, and that model will catch up.

- The randomness of the process probably allow to even go beyond just synchronizing the abilities of the models. There may be some additional complexification of quizzes that get accepted by chance.

- It can be parallelized by dispatching the GPTs across multiples nodes, and avoiding a quadratic cost by limiting the validation of the quizzes to a subset of them.

- The current process to generate new quizzes, which simply samples them at random is very rudimentary and probably not sufficient in a real-data setup. It can probably be supplemented with a MCTS-type search.

- There may be already in the generated quizzes some structure that *we* do not pick up (e.g. certain color or motion patterns).

Appendix

The code is available at

https://fleuret.org/git/culture

The experiments are done with a GTX 4090.

The GPT used has 37M parameters and the following structure:

dim_model	512
dim_keys	64
dim_hidden	2048
nb_heads	8
nb_blocks	12

Adam, $\eta = 1e-4$, no scheduling.

There are $N_{\text{train}} = 250'000$ original quizzes for training and $N_{\text{test}} = 10'000$ for test.

At each epoch, for both train and test samples, we mix original quizzes and the generated ones.

For training for instance, if there are less than $N_{\text{train}}/2$ new quizzes, we take all of them, otherwise we sample $N_{\text{train}}/2$ of them without replacement, and then we sample without replacement enough original quizzes to get N_{train} samples in total.

We proceed similarly to get N_{test} samples for test.