**General Information**

[Universal-computation](https://drive.google.com/drive/folders/1-VEiP3rj3rtYICYZSZc1enz7QVfRnG-2?usp=sharing) is our root folder. It was basically forked from <https://github.com/kzl/universal-computation>, the code from the original paper. We’ve done work in two main files

1. [Results Reproduction](https://colab.research.google.com/drive/1CJyaCzZZgD4J1esEP-b73r3PME0Y6w3U?usp=sharing): contains our experiments to reproduce the results from the paper. In the “What We’ve Done” section, we note what we were able to achieve and not achieve here.
2. [New Models, Bit Tasks](https://colab.research.google.com/drive/1ESuVSTnp8mErjzM3mhKJmUk0IGwy7Dey?usp=sharing): In this document, we started to apply new models to the simplest tasks that we were able to reproduce the results for (Bit Memory, XOR, MNIST) so that we can have some preliminary results for our draft.

The rest of the files are from the original github repo. Google Drive is a bit weird, so if you need to edit a .py file, click “Open With” and then select “Text Editor.” You’ll have to restart any colab notebooks to get them to load the new code.

[This](https://www.overleaf.com/project/6195c56864f9d05326e28cac) is the current project draft. Looks like we probably want it to be around 4 pages.

**What We’ve Done**

The original paper has 7 tasks. I reproduced accuracy standards on 3 (Bit Memory, XOR, MNIST) of them. ListOps runs fast enough, but I could not get the desired accuracy. CIFAR-10, C10 LRA, and Homology will all take on the order of hours to train.

TODOs (not in prioritized order)

1. Reproduce the results on the other 4 tasks
2. Apply new models to the 3 tasks we’ve reproduced results for
3. Apply new models to the 4 tasks we haven’t reproduced results for
4. Explore new modalities (write the code and methodology, probably don’t have time to run it)
5. Build out the sections so that this looks like an actual ACL paper. (this is the example paper from lecture--<https://arxiv.org/pdf/1606.02960.pdf>)

Next steps / order of priority:

* The next worksession should focus on 2 first. We already put together most of the experiments and ran them in [New Models, Bit Tasks](https://colab.research.google.com/drive/1ESuVSTnp8mErjzM3mhKJmUk0IGwy7Dey?usp=sharing), so **we just need to put those together into a results table**. One of the new models doesn’t work, so that needs debugging.
* The next most feasible task after that is probably sketching out some code for 4: Each downstream task gets its own file in universal\_computation/datasets. (We need to write a dataset class for each dataset.) We think the difficulty here depends on how annoying each dataset is - TIMIT (Speech recognition - audio files) and Tox21 (strings) will probably be most reasonable here (compared to ShapeNet). **We need to write the code and then a small methodology section.**
* The last thing we need to do before turning in the draft is #5. We need to **write some more math in a background section and detail very concretely what we have done/need to do in the methods section.**
* (does not need to be done before draft, but we should try it) After that, #1 is probably the next most feasible. There’s very little code to change here, since it’s hyperparameter tuning, but the bottleneck is the amount of time it takes to train + getting the right hyperparameters (which aren’t exactly specified in the original paper). Alex spent a few hours on this but still could’t get these 4 to work