

June 2015 Work Log

Jeremy Rogers
jroger44@vols.utk.edu

June, 2015

Contents

1	Goals for the Month	1
2	Progress/Notes	2
2.1	Gain ownership of the cubmisc github repository	2
2.2	Finish running the NSE first order approximation	2
2.3	Profile the test genomes	2
2.4	Read and understand the DIMCMC paper	2
2.5	Get access to Preston's repository	2
2.6	Clean up my log repository	3

1 Goals for the Month

1. Gain ownership of Logan's github repository
2. Finish running the NSE first order approximation model on all test genomes
3. Profile the test genomes
4. Read and understand the Murray et al. paper on doubly intractable MCMC.
5. Get access to Preston's repository, and begin looking through the code.
6. Clean up my log repository

2 Progress/Notes

2.1 Gain ownership of the cubmisc github repository

1. In order to gain ownership of a github repository, the current owner has to initiate the process. So I'll need to get in contact with Logan.

2.2 Finish running the NSE first order approximation

1. This is in progress. I've run the REU13 yeast, Preston's Yeast, and Logan's yeast. Now I need to run it on the other included genomes.
2. Afternoon of June 1: I have begun to run the simulation with 6000 samples on the Brewers' Yeast. Now that we have a good idea of how well the simulation fits things, we'll see how it looks on an actual genome.

2.3 Profile the test genomes

1. This is also in progress, although it may take a while to get done, since I'll have to run things numerous times.

2.4 Read and understand the DIMCMC paper

1. I have read over the paper a couple of times, but it's a pretty heavy paper, and it might take me a while to digest it fully.

2.5 Get access to Preston's repository

1. Dr. Gilchrist has informed me that he will get me access soon.
2. Update: I now have access to the repository, and I've been looking through the code. It's actually really well documented, so this shouldn't be too bad.
3. June 2: Preston's code is very well documented, and the code itself is clean. That's a pleasant surprise. There are numerous compiler warnings that I'd like to get rid of though.

4. June 2: I've gotten rid of all of the compiler warnings in the library source and the data simulation, except for one that would require some logic modification. I'll go back and fix it once I'm sure it's working before that change. I also need to fix some compiler warnings in the eta simulation code.
5. All of the compiler warnings are now fixed. There was also a bug that I found at line 259 of `seqfuncs_CES.c`. The line read

```
xi += (A1 + A2 * (i)) * (*(sigma_vec - 1) - *(sigma_vec++));
```

The problem with this is that this is undefined behavior. There is no sequence point between the `sigma_vec - 1` and the `sigma_vec++`. This means that there is no guarantee that the subtraction will take place before the increment happens. In fact, I would guess that if I looked at the assembly code, that the increment would happen first, simply because it has to move something into memory. Anyways, I changed this to

```
xi += (A1 + A2 * (i)) * (*(sigma_vec - 1) - *(sigma_vec));  
sigma_vec++;
```

So now the increment will happen after the statement guaranteed. There are comments later in the code about `sigma` doing weird things, so maybe that will fix some of them.

2.6 Clean up my log repository

1. My repository has been cleaned significantly, thanks to some scripts that I wrote. Now, I have separate repositories for logs, source files, and graphs. The logs directory only contains `.pdfs`, and the source directory only contains `.tex` files. All auxiliary files are now stored locally on my machine so that they don't clutter up the repository anymore.