Project 2 Report

Run Instructions: [# atoms] [bucket width] [threads per block]

For 10000 500.0, 32 is the optimal number of threads per block ~31-32ms

For 20000 500.0, 128 is optimal ~119ms

For 50000 500.0, 32 is optimal ~689ms

For 100000 500.0 256 is optimal ~2700ms

Techniques Used:

Privatized output:

For each block of the kernel, I create a shared memory array for the output to be written to. The global histogram must have type of long long, but the shared memory histogram will never need to be that large, as it only need contain the outputs for one block. For this reason, I keep shared memory use low with int type.

Read only cache:

Since the atom list will not be modified, I use the cudaMemcpyToSymbol function to indicate that the input can be safely cached.