# Cuckoo Hashing: P5, Progress

Group #7
Matt Johnerson, Michael Shihrer, Bradley White

November 23, 2016

The initial setup for the experiment has been more difficult than we had presumed. Finding ”good” hashing functions that map keys distinctly has been challenging because of the complexity of implementation of certain functions. Furthermore, the hash functions used in [1] did not translate well and only map to about 20% of the indexes in the array, presumably because of word size in different computer architectures.

A Python package, perfection, was discovered which can derive a perfect hash function for a set of unique integers [3]. It was tested on small sample sizes and proved to be correct, however it is quite slow when the set of integers is large, or the set of integers is randomly chosen from numbers exceeding 500,000. So, we will need to derive a large input set of integers, save it to an external file and derive the perfect hash function well before the experiment is set to be conducted.

The last hashing method from [2], what we're calling ”Bin Hashing”, seems largely dependent on the hashing functions used. Our initial test runs have reached a load factor of 0.63 – 0.84 using four hashing functions and dividing the hash table into four distinct quadrants for each hash function, but [2] advertised this setup as 0.97 load factor. Therefore, as previously stated, more research will need to be conducted into hashing functions and see if our results can become consistent with [2].

A projected timeline has been produced and can be viewed in the table on the next page which shows internal deadlines for objectives to complete the project and the amount of work completed already. Additionally, a repository has been setup on Github for version control and can be viewed here.

| Projected Timeline: | | | |
|---|---|---|---|
| **Date** | **Objectives** | **Personnel** | **Complete** |
| Nov. 22 | Implement Perfect Hashing | Matt | 50% |
| Nov. 22 | Implement Bin Method | Brad | ✓ |
| Nov. 23 | P5 Due | All | ✓ |
| Nov. 26 | Design Experiment | All | 40% |
| Nov. 26 | Implement Cuckoo Hashing | Mike | |
| Nov. 27 | Implement Experiment Driver | All | 25% |
| Nov. 28 | Code Review, Testing, & Update | All | |
| Dec. 1 | Run Experiment | All | |
| Dec. 2 | Analyze Results | All | |
| Dec. 2 | Design Video | All | |
| Dec. 4 | Produce Video | All | |
| Dec. 5 | P6 Due | All | |

Table 1: Group #7's projected timeline and completed work thus far.

# References

[1] Pagh and, Rodler. *Cuckoo Hashing*. 2001.

[2] Erlingsson, Manasse, and Mcsherry. *A Cool and Practical Alternative to Traditional Hash Tables*. 2006.

[3] `https://github.com/eddieantonio/perfection`