



**CALIFORNIA STATE SCIENCE FAIR
2009 PROJECT SUMMARY**

Name(s) Abraham P. Karplus	Project Number J1605
Project Title Find It Fast: Hash Tables and Collision Resolution	
Abstract Objectives/Goals Hash tables are an efficient method for data storage and access. They allow accessing data without extensive search. When two keys hash to the same location, a collision resolution algorithm is used to determine where the keys should be placed. I studied three collision resolution algorithms (CRAs): linear probing, linked list, and cellar. My hypothesis was that all three hash table CRAs would perform about equally well, and all much better than a simpler linear search method. Methods/Materials I measured time, memory usage, and number of string comparisons (as a more precise proxy for time) on 50 different inputs for these three hash methods and for linear search. Results As expected, hashing was a hundred times faster than linear search, taking 0.35-0.5 seconds to look up the approximately one million words of the Bible, instead of 45 seconds for linear search. Conclusions/Discussion The cellar and linked list CRAs always performed well. Contrary to my hypothesis, linear probing sometimes performed poorly when it didn't rehash until the table was full. Rehashing when the table was only 70% full allowed linear probing to perform on par with linked list and cellar.	
Summary Statement I programmed and measured the performance of three collision resolution algorithms for hash tables.	
Help Received Father taught me C, gave some help in debugging code, and helped create Makefile. Parents and teacher provided feedback on report.	