

Report On Hashing

Two hash functions were used to generate the report. Performance differences between these hash functions are not as pronounced as the difference found among different collision resolution methods, especially as the hash table size was varied.

Hash Function 1

Horner's method of hashing was used to hash strings.

Hash Function 2

FNV-1A hash function was used to hash strings.

Collision Resolution

The collision resolution methods used were:

1. Separate Chaining
2. Double Hashing
3. Custom Hashing (Double hashing and Quadratic hashing)

Results

For table sizes below 10000 double hashing and custom probing was not feasible, because 10000 unique words could not be hashed into a table of lesser size.

Since table size needs to be relatively prime with the auxiliary hash function value, table size was chosen to be the **nearest prime** of the input value.

For table size = 10,000

	Hash 1		Hash 2	
	Number of collisions	Average probes	Number of collisions	Average probes
Chaining Method	3876	1.51	3707	1.463
Double Hashing	59069	5.873	63809	6.376
Custom Probing	60160	6.349	54095	7.533

For table size = 20,000

	Hash 1		Hash 2	
	Number of collisions	Average probes	Number of collisions	Average probes
Chaining Method	2510	1.326	2057	1.247
Double Hashing	3860	1.399	3777	1.364
Custom Probing	3772	1.368	3952	1.384

For table size = 40,000

	Hash 1		Hash 2	
	Number of collisions	Average probes	Number of collisions	Average probes
Chaining Method	1381	1.156	1130	1.125
Double Hashing	1573	1.167	1540	1.147
Custom Probing	1590	1.17	1493	1.158

For table size = 60,000

	Hash 1		Hash 2	
	Number of collisions	Average probes	Number of collisions	Average probes
Chaining Method	971	1.11	789	1.09
Double Hashing	876	1.105	937	1.093
Custom Probing	865	1.1	941	1.092

For table size = 80,000

	Hash 1		Hash 2	
	Number of collisions	Average probes	Number of collisions	Average probes
Chaining Method	717	1.079	609	1.054
Double Hashing	664	1.075	703	1.079
Custom Probing	658	1.075	683	1.071

For table size = 100,000

	Hash 1		Hash 2	
	Number of collisions	Average probes	Number of collisions	Average probes
Chaining Method	560	1.054	443	1.045
Double Hashing	574	1.061	532	1.057
Custom Probing	572	1.06	554	1.06

Conclusion

From the results it is clear that as the hash table size increases, the performance of double and custom probing improves. However, when the table size nearly equals the number of keys (table size = 10000), double hashing and custom probing performs poorly. Because the density of the keys get large, and empty slots to hash into are more uncommon. However, from table size greater than or equals 20000, we see rapid improvement of the performance of double hashing and custom probing. Because the number of empty slots increases as the table size increases and therefore the number of collisions and probings required to insert and fetch an item decreases.