

Announcements:

1. HW 11 is due next Thursday (it is a very long assignment (I decided not to add more problems), so you have a week to do it.
2. B-trees (section 7.4 in Levitin should be straightforward for those who have had experience with other balanced trees, so I am asking you to read this section on your own and ask questions next week about anything you do not understand.
3. Exam 2 Tuesday Nov 4 in class
4. In my office today: hours 6, 9, 10, and maybe the last part of 8.

Main ideas from today:

1. **Clarification of rehashing** (a.k.a. double hashing). If the primary hash function $h_1(obj)$ leads to a collision at location n in the table, compute $k=h_2(obj)$ where h_2 is the secondary hash function. Then do probes at $n+k, n+2k, \dots$ until an empty spot is found (all indices are taken mod M , where M is the table size).
2. Show by contradiction that if p is a prime number larger than 3, if i and j are $\leq \lfloor p/2 \rfloor$, and if $i \neq j$, then $H + i^2 \not\equiv H + j^2 \pmod{p}$.

5. In string search algorithms, to what do the terms *pattern* and *text* refer?

pattern:

text:

6. What is the main source of inefficiency in brute force string searching?
7. What is unusual about the comparison order in Horspool's string search algorithm?
8. If the text character that we compare to the last pattern character occurs nowhere in the pattern, how far right can we shift the pattern, and still be sure that we do not miss a match?
9. Explain the reasoning behind the rule for calculating $t[c]$ in the Horspool shift table.

10. Show the values in the Horspool shift table for the following characters if the pattern is COCACOLA :

Character	Shift
C	
O	
A	
L	
M	

