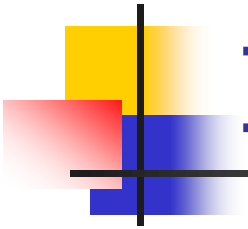
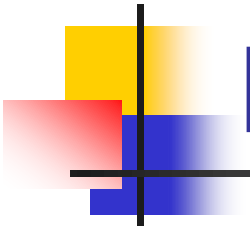


Internal Sort Algorithm

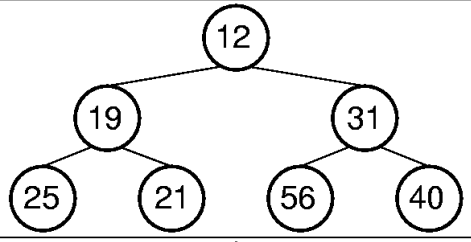
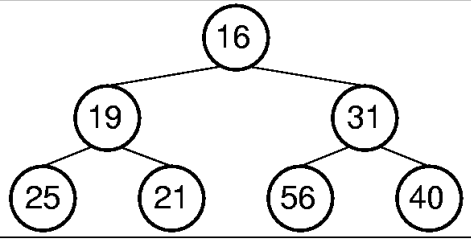
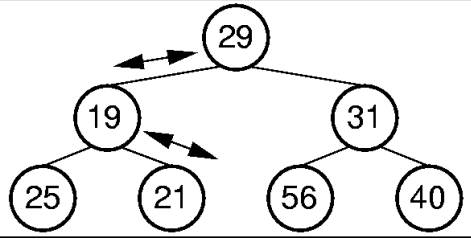
- Quicksort is a fast way to sort in memory
- Alternative: “tournament sort” (a.k.a. “heapsort”, “replacement selection”)
  - Break available memory into an array for the heap, an input buffer, and an output buffer
  - Fill the array from disk
  - Make a min-heap
  - Send the smallest value (root) to the output buffer

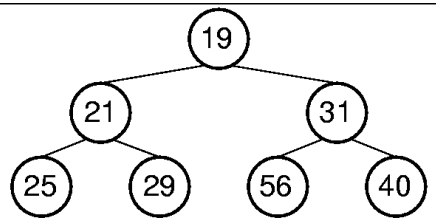
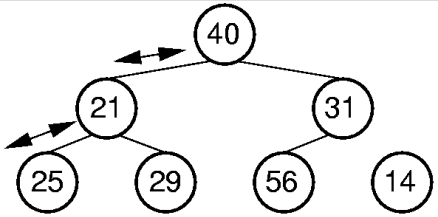
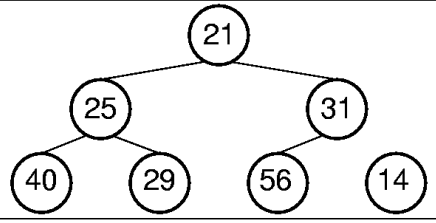


Replacement Selection

- If the next key in the file is greater than the last value output, then
  - Replace the root with this key
- else
  - Replace the root with the last key in the array
  - Add the next record in the file to a new heap (actually, stick it at the end of the array)

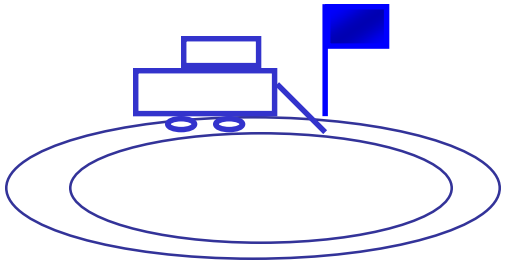
An Illustration of Replacement Selection

Input	Memory	Output
16		12
29		16
		

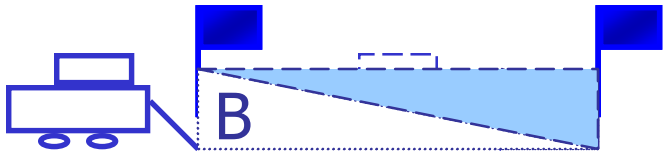
14		19
		
35		21

More on Heapsort

- Fact: average length of a run in heapsort is *2(B-2)*
  - The "snowplow" analogy



- Worst-Case:
  - What is min length of a run?
  - How does this arise?
- Best-Case:
  - What is max length of a run?
  - How does this arise?



- Quicksort is faster, but ... longer runs often means fewer passes!