Weekplan: External Memory

Philip Bille

References and Reading

[1] Cache-oblivious dynamic programming, R. A. Chowdhury and V. Ramachandran, SODA 2006.

[2] The Input/Output Complexity of Sorting and Related Problems, A. Aggarwal and J. Vitter, CACM 1988

[3] Cache-Oblivious Algorithms and Data Structures, Erik Demaine, Lecture Notes from the EEF Summer School on Massive Data Sets, 2002

We recommend reading [1] and [3] in detail. [3] define the I/O model and cache-oblivious model and covers most basic cache-oblivious algorithms. [1] presents the cache-oblivious algorithm covered in the lecture. [2] is the original paper defining the I/O model.

Exercises

1 String Reversal Let S be a string of length N stored in $O(N/B)$ blocks. We want to compute the *reverse* string S^R of S. Solve the following exercises.

1.1 Give an efficient algorithm to reverse S in the I/O model.

1.2 Give an efficient algorithm to reverse S in the cache-oblivious model.

2 Stacks and Queues in External Memory Show how to implement stacks and queues with $O(1/B)$ amortized I/Os per operation in the I/O model of computation.

3 External Sorting We want to sort an array of N numbers in the I/O model efficiently. Solve the following exercises.

3.1 Show how to merge $\Theta(M/B)$ sorted arrays of total length N into a single sorted array in $O(N/B)$ I/Os.

3.2 Given an unsorted array of length N, show how to create $\Theta(N/M)$ sorted arrays of each of length M in $O(N/B)$ I/Os.

3.3 Show how to sort an array of length N using

$$O\left(\frac{N}{B}\log_{M/B}\frac{N}{M}\right)$$

I/Os. *Hint:* Do a multiway merge using 1 for merging and 2 as base case.

4 Parallel Dynamic Programming Consider the standard dynamic programming algorithm for the shortest path in implicit graphs problem. Suppose that we have $p > 1$ processors at our disposal. How can we use these to speedup the standard dynamic programming solution in the RAM model? What about the I/O model or cache-oblivious model?

5 Dynamic Programming meets Divide and Conquer Consider the standard dynamic programming algorithm for the shortest path in implicit graphs problem on a RAM model of computation. We are interested in efficiently computing not only the length of the shortest path but also the edges on the shortest path. Solve the following exercises.

5.1 Show that with $O(n^2)$ space we can compute the path in $O(n^2)$ time.

5.2 Show that we can compute a single edge on the shortest path corresponding to the $n/2$th row in the graph using $O(n)$ space and $O(n^2)$ time.

5.3 Show how to recursively apply 2 to output the shortest path in $O(n)$ space and $O(n^2)$ time.

6 Medians Let A be an array of N numbers. Show how to find the median of A in $O(N/B)$ time. *Hint:* The classical divide and conquer linear time RAM algorithms works.