Math 262A: Circuit Complexity

AKS Sorting network

Lecturer: Sam Buss
Scribe: Radheshyam Balasundaram

December 20, 2013

In previous class we saw an example of oblivious sorting network, Batcher Sort which had a depth of $O((\log n)^2)$. In this lecture, we will see a $O(\log n)$ depth cirsuit for the same. This circuit was first given by Ajtai, Komlós and Szemerédi in 1983 (see [2]). It was simplified by Paterson (see [3]) and later by Seiferas (see [4]). In this lecture we will see construciton by Seiferas as that is the easiest to describe.

One of the main components of the network is $\epsilon-$Approximate $\lambda-$Camparator/ $\epsilon-$Halver. Let us recall it's definition from previous class:

Definition 1. Let $\epsilon > 0$ and $0 \leq \lambda \leq \frac{1}{2}$. A sorting network is an $\epsilon-$*Approximate* $\lambda-$*Camparator/* $\epsilon-$*Halver* on n inputs iff for all $\lambda' \leq \lambda$ at most $\epsilon\lambda' n$ of the $\lfloor \lambda' n \rfloor$ many smallest inputs are not among the leftmost $\lfloor \lambda n \rfloor$ many outputs. Dually, at most $\epsilon\lambda' n$ of the $\lfloor \lambda' n \rfloor$ many largest inputs are not among the rightmost $\lfloor \lambda n \rfloor$ many outputs.

We call an ϵ-approximate $\frac{1}{2}$-comparator an ϵ-*halver*.

Theorem 2. *Let $\epsilon > 0, \lambda \leq \frac{1}{2}$. There is a $d = d(\epsilon, \lambda)$ such that $\forall n = 2m$, there is a sorting network of depth d which is both $\epsilon-$Approximate $\lambda-$Camparator/ $\epsilon-$Halver.*

We shall assume the existance of these approximate sorting networks for now and use them in the AKS sorting network. We shall prove the above theorem later.

The AKS sorting network

Theorem 3 ([2])**.** *There are sorting networks of depth $O(\log n)$.*

Corollary 4. *There exist monotone formulas over bounded fanin $\cap, \vee, \neg$ gates of depth $O(\log n)$ for Threshold function, Th_k^n. Hence, they are of size $n^{O(1)}$.*

Proof of AKS sorting by [4]. Without loss of generality, assume n is a power of 2, $n = 2^l$. Also, assume all the inputs are distinct.

Components of the network

The sorting network has *levels* and each level has n *wires* or *signals* each of which holds one of the input values.

At any level, these wires will be put into *bags*. There are $n-1$ bags, each bag represents a binary sub-interval of $0, 1, \ldots n-1$. There is one bag for the whole interval. There are two bags for left and right half-intervals and 4 bags for quarter intervals and so on until $n/2$ bags for contiguous pairs. Total number of bags is $1 + 2 + 4 + \ldots + n/2 = n - 1$.

These bags can be imagined as nodes of binary tree with the root node (with *depth* 0) being the first bag for whole interval and each bag b (at depth d) corresponding to an interval (of length $n/2^d$) has two *children bags*, b_l, b_r, corresponding to left and right sub-intervals. b is said to be *parent* of b_l, b_r and b's parent (if one exists) is called a *grandparent* of b_l, b_r. To begin with, the input to sorting network belong to depth 0 bag.

A signal is said to be *native* to its bag if the rank value of the signal (on the wire) is a member of the interval associated with the bag. Otherwise, the signal is called a *stranger*.

A wire is a *j-stranger* if the distance from the bag holding the signal to the closest bag it is native to in the bag-tree hierarchy is at least j. Hence, a j-stranger is also a k-stranger for $1 \leq k \leq j$. 1-strager is same as stranger and all signals are 0-strangers.

With each bag at depth d during round t, we associate a capacity value $b = b(d, t)$ which will bound the number of signals in the bag. Note that this is different from the length of the interval associated with the bag.

Parameters in the network

We define a few parameters which we shall use in the construction of network: $\epsilon = \lambda = \frac{1}{99}, \nu = 0.65, A = 10$.

The capacity function is given by:

$$b(d, t) = n\nu^t A^d \tag{1}$$

The capacity increases with depth d and decreases with number of rounds t.

Rounds of sorting

Initially all signals are in root bag. After t rounds, do the following in $t + 1^{th}$ round. Conisder a non-empty bag with N signals in it. Let b be its capacity, $b \geq N$.

1. Apply $\epsilon-$Approximate $\lambda-$Camparator/ $\epsilon-$Halverto these N elements.
2. If not the root bag, some elements may not be native to this bag and need to be pushed to parent bag. Hence, we move leftmost $\lfloor \lambda b \rfloor$ and rightmost $\lfloor \lambda b \rfloor$ to the parents. If N is odd, we randomly select an element and move it to parent
3. Of the remaining signals, the leftmost half of wires go to the left child and dually for the right.

We stop this process when the lowest level (leaf level) bags have capacity $b < 1/\lambda$.

Properties of bags

Now, we prove some observations regarding the bags:

Property 1: At any round, alternative levels are emtpy. The symmetry of bags at a fixed level ensures that all bags at a particular level have same number of signals.

Property 2: Number of signals in a subtree of a bag rooted at $b \leq$ length of the interval associated with b

Property 3: Number of signals in a bag $b \leq$ capacity of b.

Property 4: Number of $j-$strangers $\leq \lambda\epsilon^{j-1}b(d,t)$.

Property 5: Leaf bags (that correspond to intervals of length 2) do not push the elements to the bottom.

Now, revisiting the stopping condition, when leaf level has capacity $b < 1/\lambda$, using property 1, its parent is empty. And, capacity of it's grandparent is $bA^{-2} = 1/\lambda A^2 < 1$ (from eqn 1).

Also, when this happens, observe that from Property 4, bags have no stragers. Number of 1-strangers $\leq \lambda\epsilon^{1-1}b(d,t) = \lambda b(d,t) < 1$.

Now we prove the properties listed above. Property 1 is straightforward from the construction.

Proof of Property 5 If we have not reached stopping condition, capacity of the leaf node is $\geq 1/\lambda$. Of these, we send up $2\lambda \cdot 1/\lambda$ of them.

Proof of Property 3 We prove this using induction on t.

Capacity at depth 0 is $n\nu^0A^0 = n \geq n$. For induction, assume that the statement is true for all $t' \leq t$.

Consider a non-empty bag after $t+1$ rounds. We want to show that the statement is true for this bag. Let the capacity of this bag at round t be $b = n\nu^t A^d$. In next round, its capacity function will be $b\nu$. We shall prove that the number of signals does not exceed this capacity at round $t+1$.

After t rounds, by induction hypothesis, parents have $\leq b/A$ signals and children $\leq bA$ signals each. Parents send down $\leq 1/2b/A$ signals down and children each send up $\leq (2\lfloor\lambda bA\rfloor + 1)$. So, total number of elements is:

$$\leq \frac{b}{2A} + 4\lfloor\lambda bA\rfloor + 2 \tag{2}$$

Now, if $b \geq A$, the above equation 2 becomes:

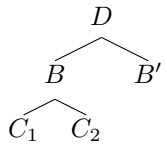
$$\leq \frac{b}{2A} + 4\lfloor\lambda bA\rfloor + 2 \leq b(\frac{1}{2A} + 4\lambda A + \frac{2}{A}) \leq b \cdot \nu$$

If $b < A$, grandparent's capacity is $b/A^2 < 1$ and parent's capacity is $b/A < 1$. Hence the children of this bag have even number of signals.

Thus number of signals sent to the bag from children at $t + 1^{th}$ round is $\leq 4\lfloor\lambda bA\rfloor < b \cdot \nu$.

Proof of Property 4 Again, we use induction on t. For $t = 0$, there are no strangers. Now, assume it is true til t.

Case: If $j > 1$. Let B be a bag with parent bag D and children C_1, C_2 and sibling B'.



At $t = 0$, B is empty. We want to bound the number of strangers at time $t + 1$. Let b be the capacity of b at time t. Where can the $j-$strangers come from?

1. $j - 1$ strangers of D that get sent down to B. D has $\leq \lambda\epsilon^{j-2}(b/A)$ many $j - 1$ strangers. Due to the $\epsilon-$approximation property, at most $< 2\epsilon$ fraction of these $j - 1$ strangers get sent down, which is $< 2\lambda\epsilon^{j-1}\frac{b}{A}$.

2. $j + 1$ strangers in C_1 and C_2 get sent up to B. There are $\leq 2\lambda\epsilon^j bA$ many $j + 1-$strangers.

So, total number of $j-$strangers is $\leq \lambda\epsilon^{j-1}\frac{b}{A} + 2\lambda\epsilon^j bA$. For our proof, want this to be $\leq \lambda\epsilon^{j-1}\nu b$. This is true because by choice of our parameters, $\nu \geq (\frac{2}{A} + 2\epsilon A)$.

Now, we shall prove this property for the case when $j = 1$. Let B, D, C_1, C_2, B' be the same as defined before. Here, 1-strangers at B at time $t + 1$ can come from:

1. 1-stranger in D, which is $\leq 2\lambda b/A$.

2. 2-stranger in C_1, C_2, which is $\leq 2\lambda\epsilon bA$

3. Items in D that are native to D and B' bug erraneously sent to B. This can be split into two cases:

   (a) $\epsilon-$halving errors: $\leq \epsilon(1/2\text{capacity of } D) \leq \frac{\epsilon b}{2A}$.

   (b) There could me more B' natives in D than B natives (instead of equal numbers). We shall count this now.

   Compare the position of signals in bags to a *reference distribution.* In this reference distribution, (i) if B" is a bag in the same level of B, all members of bags below B" are native to B". (ii) Any bag D' at the level of D is holding number of elements native to its two children.

   Number of B' natives above level D in reference distribution is, due to symmetry of bags, $\leq 2^{-d}[\frac{b}{A^3}2^{d-3} + \frac{b}{A^5}2^{d-5} + \ldots]$ where 2^{-d} is the number of bags in level d and bA^3 is the capacity of grandparent of D (bag that is at level $d-3$), 2^{d-3} is the number of bags at the same level and so on. This evaluates to $\leq \frac{b}{2A}\frac{1}{4A^2-1}$.

   Number of strangers in or below B' bounds number of B' native elements from the reference distribution that can be moved up to D. This is bounded by:

$$\leq \Sigma(bA^k) \cdot 2^k \cdot \lambda\epsilon^{k+1-1} \leq \frac{2\lambda\epsilon bA}{1 - (2\epsilon A)^2}$$

where bA^k is the capacity of decendant at level k from B' and 2^k is the number of bags at level k from B' and $\lambda\epsilon^{k+1-1}$ is the number of $k+1-$strangers at this bag.

Existance of $\epsilon-$Approximate $\lambda-$Camparator/ $\epsilon-$Halver

Now, we shall show existance of $\epsilon-$Approximate $\lambda-$Camparator/ $\epsilon-$Halver. This completes the proof of AKS sorting network. For this, we use the existence of *Expander Graphs* defined below.

Lemma 5. *There exist $d-$regular bipartite graphs G of $2n$ vertices such that $\forall S$, subset of one of the parts of vertices, $|N(S)| \geq \min\{|S| \cdot \frac{(1-\epsilon)}{\epsilon}, (1-\epsilon)n\}$, where $N(S)$ is set of neighbors of S. Further more, this graph is a union of $d-$many perfect matchings.*

Lemma 6. *$\epsilon-$ halvers exist. That is, let $\epsilon > 0$. $\exists d > 0$ such that $\forall$ even n, $\exists\epsilon-$halver of depth d.*

Proof. We use the expander graph on $n = 2m$ vertices and do d rounds of comparions using m nodes of the leftmost signals and m of the rightmost signals. Now, let us prove that this is an $\epsilon-$halver.

Let S be the set of leftmost outputs of elements in the wrong half. Suppose $|S| > \epsilon m$, $|N(S)| \geq |S| \cdot \frac{(1-\epsilon)}{\epsilon} \geq (1-\epsilon)m$. So, $|N(S)| + |S| > m$. This is a contradiction because S along with it's neighbors should contain elements from smaller half. □

Lemma 7. *Let $\epsilon > 0$ and $0 < \lambda < 1$. There is a $d > 0$ and a $d-$depth sorting network which is both $\epsilon-$Approximate $\lambda-$Camparator/ $\epsilon-$Halver.*

Proof. We are going to repeatedly apply ϵ_0-halver for some $\epsilon_0 < \epsilon$ (to be defined later).

First apply an ϵ_0-halver on n elements. Now, take $m = \lfloor \lambda n \rfloor$. Repeatedly use ϵ_0-halvers for intervals of size $2m, 4m, 8, \ldots, 2^{\lceil \log \frac{n}{m} \rceil - 1}$ in reverse order of the list (but not sitching out original halves).

Let $\lambda' \leq \lambda$. Consider the $\lfloor \lambda' n \rfloor$ many least elements. We want at most $\epsilon\lambda' n$ of these not in the bottom$\lfloor \lambda n \rfloor$. This happens because of ϵ_0-halver misconfigurations. Each ϵ_0-halver misconfigures at most $\epsilon_0\lambda' n$ items and total is $\leq \epsilon_0 \lceil \log \frac{n}{m} \rceil \lambda' n$. So, we choose a value for ϵ_0 such that $\leq \epsilon_0 \lceil \log \frac{n}{m} \rceil \leq \epsilon$. □

This concludes the proof of Theorem 3. □

References

[1] Kenneth E Batcher, *Sorting networks and their applications*. Proceedings of the April 30–May 2, 1968, spring joint computer conference, 1968.

[2] Miklor Ajtai, Jáno Komlós, Endre Szemerédi, *An $O(n \log n)$ sorting network*. STOC, 1983.

[3] MS Paterson, *Improved sorting networks with $O(\log n)$ depth*. Algorithmica, 1990.

[4] J Seiferas, *Sorting networks of logarithmic depth, further simplified.* Algorithmica, 2009.