Recitation 1

Introduction

1.1 Announcements

- Welcome to 15-210!
- The course website is http://www.cs.cmu.edu/~15210/. It contains the syllabus, schedule, library documentation, staff contact information, and other useful resources.
- We will be using Piazza (https://piazza.com/) as a hub for course announcements and general questions pertaining to the course. Please check it frequently to make sure you don’t miss anything.
- The office hours schedule is posted on the course website as well as Piazza. Come meet all of your TAs!
- The first homework assignment, IntegralLab, has been released! It’s due Friday at 5pm, but don’t worry – it’s quite short.
- Homeworaks will be distributed through Autolab (https://autolab.andrew.cmu.edu/). Most homework assignments will be released on Fridays and will be due one week later. You will submit coding tasks on Autolab, and written tasks on Gradescope (https://gradescope.com/).
1.2 Sequences

Sequences, which you’ve seen in 15-150, are a parallel data structure that are immutable and functional like lists, yet easy to randomly access like arrays. They’re a useful abstraction for various real-world parallel sequence structures that let us model parallel computation at the algorithmic level.

Several basic operations are useful on sequences, like tabulations, maps, filters, and reductions. We use mathematical notation to represent these manipulations:

\[
\begin{align*}
\langle e : 0 \leq i < n \rangle & \quad \text{tabulate (fn i => e) n} \\
\langle e : p \in S \rangle & \quad \text{map (fn p => e) S} \\
\langle p \in S | e \rangle & \quad \text{filter (fn p => e) S}
\end{align*}
\]

Using an efficient sequence, like ArraySequence, the work of tabulate, map, filter, and reduce is linear in the size of the sequence if the function argument is constant-time. However, we get large gains in span: constant for tabulate and map, and logarithmic for filter and reduce.

1.3 Primes

Let’s now get some practice using parallel sequences.

Task 1.1. Implement the function \( \text{isPrime} : \text{int} \rightarrow \text{bool} \), which returns true iff the input is a prime number.

Algorithm 1.2. Primality testing.

```ml
fun isPrime n = 
  let val all = \( n \mod i : 1 \leq i \leq \lfloor \sqrt{n} \rfloor \) 
  val divisors = \( x : x \in all \mid x = 0 \) 
  in |divisors| = 1 
  end
```

Task 1.3. Determine the work and span of \( \text{isPrime} \ n \).
The work is $O(\sqrt{n})$, the size of the largest sequence, and the span is $O(\log n)$ due to the filter step. Notice that $n$ is the number, not the length of its encoding.

Now that we have a way of testing primality, we can try to generate a sequence of primes.

**Task 1.4.** Implement the function $\text{primes} : \text{int} \to \text{int seq}$, which returns a sequence of the prime numbers less than $n$.

**Algorithm 1.5.** Prime generation.

```plaintext
fun primes n = 
let val all = ⟨i : 1 < i < n⟩ 
in ⟨x : x ∈ all | isPrime x⟩ 
end
```

The work and span of this algorithm are more subtle. Using tools that we’ll develop later in the course, we are able to derive a $O(n^{3/2})$ work and $O(\log n)$ span bound for the simplest prime generation algorithm.
1.4 Bonus Exercises

There’s a more efficient way to generate primes up to $n$. Knowing a number is composite, we know that all multiples of it up to $n$ are composite. So, for each number from 1 to $\sqrt{n}$, we can keep track of its multiples, and for each composite we can mark its multiples composite. This algorithm is called the sieve of Eratosthenes, and it is much more efficient.

**Task 1.6.** Implement the function $\text{primes} : \text{int} \rightarrow \text{int seq}$ using the improved method.

**Algorithm 1.7.** Prime generation using the sieve method.

```ocaml
fun primes n =  
  let
    val cs = \langle i * j : 1 \leq i \leq \lfloor \sqrt{n} \rfloor, 2 \leq j \leq n/i \rangle
    val sieve = \langle (x, false) : x \in cs \rangle
    val all = \langle true : 0 \leq i < n \rangle
    val isPrime = inject all sieve
  in
    \langle i : 2 \leq i < n | isPrime[i] \rangle
  end
```

It turns out that the sieve algorithm improves the work to $O(n \log n)$ while span remains $O(\log n)$. A derivation of the work and span may be found in the textbook, or you can ask your TA’s about it after recitation.