Exercise 1.24. Modify the `timed-prime-test` procedure of exercise 1.22 to use `fast-prime?`(the Fermat method), and test each of the 12 primes you found in that exercise. Since the Fermat test has $\Theta(\log n)$ growth, how would you expect the time to test primes near 1,000,000 to compare with the time needed to test primes near 1000? Do your data bear this out? Can you explain any discrepancy you find?

Answer. Note that test of primality is actually performed in procedure `start-prime-test`, which is invoked by `timed-prime-test`. Thus, we can modify the `timed-prime-test` procedure by only altering the implementation of `start-prime-test`:

```
(define (start-prime-test n start-time)
  (if (fast-prime? n 1000)
      (repot-prime (- (runtime) start-time))))
```

As is discribed in the problem: the Fermat test has $\Theta(\log n)$ growth. Note that, $1{,}000{,}000 = 1000^2$, thus,

$$\begin{aligned} \Theta(\log n^2) &= \Theta(2\log n) \\ &= 2\,\Theta(\log n) \end{aligned}$$

So we expect the time to test primes near 1,000,000 would be twice to compare with the time needed to test primes near 1000.

Now, we have to perform our tests in order to verify our hypothesis. As you can see in Table 1, the test result

Magnitude	Prime	Time (s)	Average Time (s)	Ratio
10^3	1009	3.0000000000001137e-2	3.0000000000001137e-2	–
	1013	3.0000000000001137e-2		
	1019	3.0000000000001137e-2		
10^6	1000000007	0.0799999999999983	0.0766666666666656	2.55555555555542
	1000000009	0.07000000000000028		
	1000000021	0.0799999999999983		

Table 1. Time Required to Find the First 3 Prime Number in Different Magnitude

violates our prediction surprisingly. The ratio of time consumed between 1,000,000 and 1000 in testing primes turns out to be about 2.56 rather than twice, which we once believed in. But why?

Well, remember that the `time-prime-test` we implemented here uses `fast-prime?`(the Fermat method), which is a probabilistic method. In validaing the primes out of a goup of integers, the testing data picked randomly by the evaluator are surely diverse from one another. Thus, the amount of computation it involves will undoubtfully varies among different testing rounds. This reveals the cause of the discrepancy we found above.