

Homework 2 Problem 5

Programming Problem 5 (10 points):

The *Coupon Collector* problem is: There are n types of coupons. Each time you get a coupon, you are given a coupon of a random type (with equal probability of receiving each coupon). The question is how many coupons do you expect to receive, on the average, before you have collected the full set of coupons.

Your programming assignment is to write a simulator of the Coupon Collector Problem, and run simulations to see how long how many coupons are needed to complete the set. You should run your program for values of n up to 4,000. Determine the average number of coupons required to complete the set. How does this relate to the results of programming problem 5 of homework 1? (Consider both the total number of coupons C , as well as C/n .)

Answer:

```
import java.util.Arrays;

class couponCollector{
    public static void main(String [] args){
        for(int i = 500; i <= 5000; i += 500){
            float localSum = 0;
            for(int j = 0; j < 50; j ++){
                localSum += sim(i);
            }
            System.out.printf(
                "n:%d, \n"+
                "avgC:%.2f, \n"+
                "C/n:%.2f\n",
                i,
                localSum/50,
                (localSum/50)/i);
        }
    }

    public static float sim(int n) {
        int [] owned = new int[n];
        int count = 0;
```

```

        while (Arrays.stream(owned).sum() != n){
            int idx = (int)(Math.random() * n);
            if (owned[idx] == 0) owned[idx]++;
            count++;
        }
        return count;
    }
}

```

The output of coupon collector:

```

n:500, avgC:3388.88, C/n:6.78
n:1000, avgC:7539.82, C/n:7.54
n:1500, avgC:12122.86, C/n:8.08
n:2000, avgC:16065.76, C/n:8.03
n:2500, avgC:21015.84, C/n:8.41
n:3000, avgC:25416.60, C/n:8.47
n:3500, avgC:30692.40, C/n:8.77
n:4000, avgC:34984.60, C/n:8.75
n:4500, avgC:39890.24, C/n:8.86
n:5000, avgC:45706.40, C/n:9.14

```

The goodness of m in HW1P5:

```

n:500, M:3065, M/n:6
n:1000, M:9036, M/n:9
n:1500, M:14274, M/n:9
n:2000, M:17396, M/n:8
n:2500, M:22651, M/n:9
n:3000, M:28297, M/n:9
n:3500, M:32899, M/n:9
n:4000, M:28437, M/n:7
n:4500, M:35469, M/n:7
n:5000, M:49494, M/n:9

```

It is easy to spot that the total number of coupons to collected the full set of coupons and goodness of m shared a similar logarithmic function with n, with similar growth function close to $n \ln n$.