Select a random number from stream, with O(1) space

Given a stream of numbers, generate a random number from the stream. You are allowed to use only O(1) space and the input is in the form of stream, so can't store the previously seen numbers.

So how do we generate a random number from the whole stream such that the probability of picking any number is 1/n. with O(1) extra space? This problem is a variation of Reservoir Sampling. Here the value of k is 1.

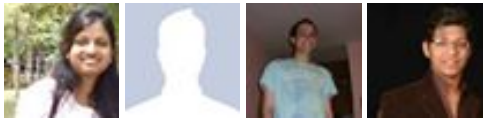
1) Initialize 'count' as 0, 'count' is used to store count of numbers seen so far in stream.
2) For each number 'x' from stream, do following
.....**a)** Increment 'count' by 1.
.....**b)** If count is 1, set result as x, and return result.
.....**c)** Generate a random number from 0 to 'count-1′. Let the generated random number be i.
.....**d)** If i is equal to 'count – 1′, update the result as x.

```
// An efficient program to randomly select a number from stream of numl
#include <stdio.h>
#include <stdlib.h>
#include <time.h>

// A function to randomly select a item from stream[0], stream[1], ..
int selectRandom(int x)
{
    static int res;     // The resultant random number
    static int count = 0;  //Count of numbers visited so far in stream

    count++;  // increment count of numbers seen so far

    // If this is the first element from stream, return it
    if (count == 1)
```

```
        res = x;
    else
    {
        // Generate a random number from 0 to count - 1
        int i = rand() % count;

        // Replace the prev random number with new number with 1/count
        if (i == count - 1)
            res  = x;
    }
    return res;
}

// Driver program to test above function.
int main()
{
    int stream[] = {1, 2, 3, 4};
    int n = sizeof(stream)/sizeof(stream[0]);

    // Use a different seed value for every run.
    srand(time(NULL));
    for (int i = 0; i < n; ++i)
        printf("Random number from first %d numbers is %d \n",
                                  i+1, selectRandom(stream[i]));
    return 0;
}
```

Output:

```
Random number from first 1 numbers is 1
Random number from first 2 numbers is 1
Random number from first 3 numbers is 3
Random number from first 4 numbers is 4
```

Auxiliary Space: O(1)

How does this work

We need to prove that every element is picked with 1/n probability where n is the number of items seen so far. For every new stream item x, we pick a random number from 0 to 'count -1', if the picked number is 'count-1', we replace the previous result with x.

To simplify proof, let us first consider the last element, the last element replaces the previously stored result with 1/n probability. So probability of getting last element as result is 1/n.

Popular Posts

Let us now talk about second last element. When second last element processed first time, the probability that it replaced the previous result is 1/(n-1). The probability that previous result stays when nth item is considered is (n-1)/n. So probability that the second last element is picked in last iteration is [1/(n-1)] * [(n-1)/n] which is 1/n.

Similarly, we can prove for third last element and others.

References:
Reservoir Sampling

Please write comments if you find anything incorrect, or you want to share more information about the topic discussed above.

Related Tpoics:

- Backtracking | Set 8 (Solving Cryptarithmetic Puzzles)
- Tail Recursion
- Find if two rectangles overlap
- Analysis of Algorithm | Set 4 (Solving Recurrences)

- Print all possible paths from top left to bottom right of a mXn matrix
- Generate all unique partitions of an integer
- Russian Peasant Multiplication
- Closest Pair of Points | O(nlogn) Implementation

Writing code in comment? Please use **ideone.com** and share the link here.

7 Comments GeeksforGeeks

Sort by Newest

Surabhi • 9 months ago

I am kind of confused with the solution you proposed everytime we are holding new number, int i = rand() % count; line suggest that if a number is generated the already stored res number nor the latest count-1 than also u r returning res equiprobable,everytime new number i is having the probability of 1/2 of selectir

abhishek08aug • a year ago

Cormen problem: 5.1-2:

```
 #include <iostream>
#include <ctime>
using namespace std;

int random(int a, int b) {
  static int first_run=1;
  if(a==0 && b==1) {
```


Recent Comments

Abhi You live US or India?
Google (Mountain View) interview · 12 minutes ago

Aman Hi, Why arent we checking for conditions...
Write a C program to Delete a Tree. · 52 minutes ago

kzs please provide solution for the problem...
Backtracking | Set 2 (Rat in a Maze) · 56 minutes ago

Sanjay Agarwal bool tree::Root_to_leaf_path_given_sum(tree...
Root to leaf path sum equal to a given number · 1 hour ago

GOPI GOPINATH @admin Highlight this sentence "We can easily...
Count trailing zeroes in factorial of a number · 1 hour ago

newCoder3006 If the array contains negative numbers also. We...
Find subarray with given sum · 1 hour ago

```
    srand( (unsigned)time( NULL ) );
    first_run=0;
  }
  return rand()%2;
} else if(a==b){
  return a;
} else if(b==a+1) {
  if(random(0, 1)) {
```

see more

Hemanth • a year ago

It's a google interview question with little variation.

Select a random quote from a given input file. Each quote can be of any no. of

Ex input file:
Quote1 Line1
Quote1 Line2
Quote1 Line3
%%
Quote2 Line1
Quote2 Line2
Quote2 Line3
Quote2 Line4
Quote2 Line5
%%
Quote3 Line1
%%
Quote4 Line1
Quote4 Line2

Quote4 Line3
Quote4 Line4
%%
Quote5 Line1
Quote5 Line2
%%

neo • 2 years ago

this code has flaw, which is, it is not selecting random number each time with number of n say 10, then for each selection it will give most latest number not

say in above example it will never give output as:

Random number from first 1 numbers is 1
Random number from first 2 numbers is 1
Random number from first 3 numbers is 3
Random number from first 4 numbers is 2

kartik neo • 2 years ago

@neo: please note that the question is not about selecting a set of ran random number at any point in stream. You will never get 2 after 3, but may get 2 in the 4th iteration with 1/n probability.

pradeep gupta • 2 years ago

nice solution.

```
/* Paste your code here (You may delete these lines if not writing c
```

Lokesh • 2 years ago

awesome, it's a google interview question