**INFO 6205**

**Program Structures & Algorithms**

**Summer Full 2018**

**Assignment 2**

Union find is to determine whether the elements are in the same data set and merge the disjoint elements into existing sets until there is only one data set(component) remain.

In this assignment, the number of sites are various integers from 1-2000 and I take a test number from the command line to deduce expression out of them. The specific experiment steps are described below.

1. **CONCLUSION:**

The number of pairs generated to accomplish n-sites union (i.e. to reduce the number of components from n to 1) is *~ 1/2 n ln n* where *ln n* is the natural logarithm of *n.*

I have deduced to an expression given below:

p=1/2 \* n \* ln(n)

Some useful abbreviations:

* p - The number of pairs generated to accomplish n-sites union
* n – The numbers of sites

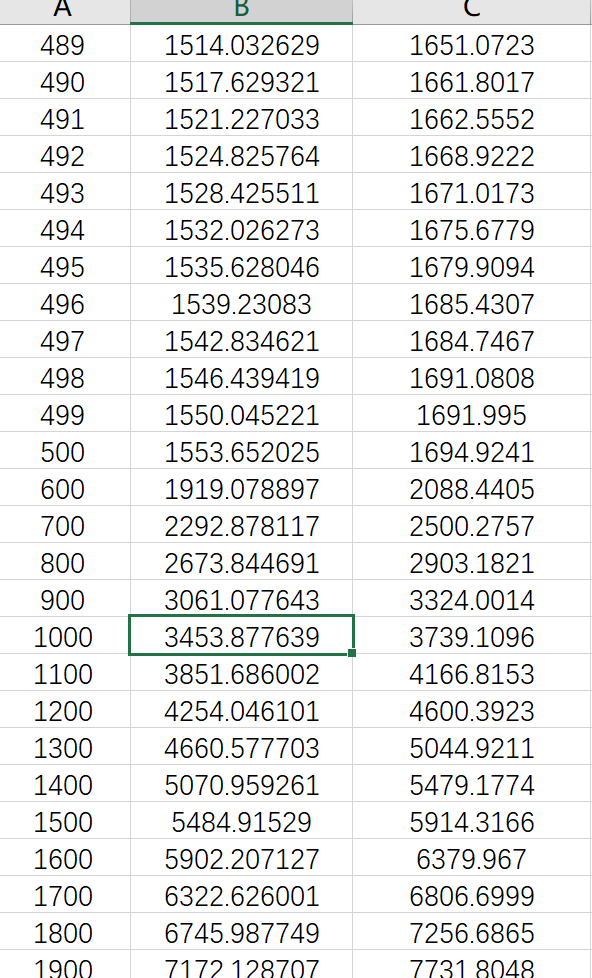
In this program, various sites from integer 2 to 2000 are been tested. For the sites interval 2-500, I ran every integer number like n=2, n=3, …, n=499; and for the sites interval 500-2000, I ran every 100 number like n=500,n=600, … , n=200.

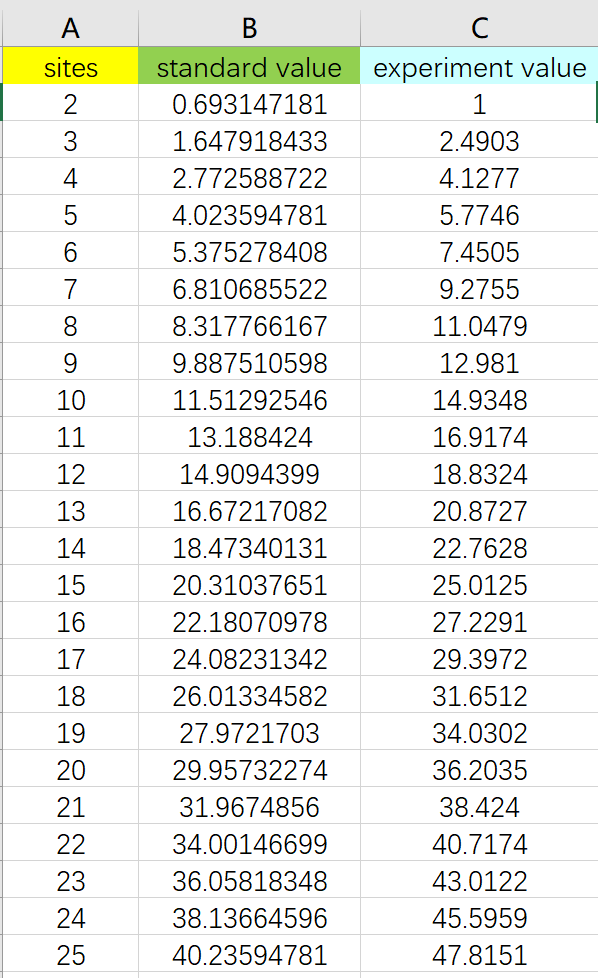
As for each single site, I set t=10000.It means the union-find operation of this site will be ran 10000 times repeatedly. And finally take the average of p. This procedure can be deduced into the expression :p/ 10000.

Total experiments times: (498+16) \*10000=5140000

1. **GRAPHS of various sites numbers**

Parts of the data I got show in the screen shot below:





The entire data set can been seen in the attachment file.

Let’s take an example to prove the equation :

(standard value is calculated from the equation, it is a theoretical value)

(test value is from the experiments )

**Example1:**

n=200

standard value: p=1/2\*200\*ln(200)=589.8317

test value: 589.59

Although these two values have some variation, both are almost same. So it is equal.

**Example2:**

n=500

standard value: p=1/2\*5000\*ln(500)=1553.6520

test value: 1694.924

Although these two values have some variation, both are almost same. So it is equal.

**Example2:**

n=1000

standard value: p=1/2\*1000\*ln(1000)=3453.8776

test value: 3739.11

Although these two values have some variation, both are almost same. So it is equal.

**Example2:**

n=1800

standard value: p=1/2\*1800\*ln(1800)=6745.9877

test value: 7256.687

Although these two values have some variation, both are almost same. So it is equal.

**HENCE PROVED**