CODEFORCES
Sponsored by Telegram

| 🇬🇧 🇷🇺
nathan_luiz | Logout

HOME   TOP   CATALOG   CONTESTS   GYM   PROBLEMSET   GROUPS   RATING   EDU   API   CALENDAR   HELP

NISIYAMA_SUZUNE   BLOG   TEAMS   SUBMISSIONS   GROUPS   CONTESTS

Nisiyama_Suzune's blog

[Tutorial] Math note — Möbius inversion

By **Nisiyama_Suzune**, 5 years ago, 🇬🇧

If you've ever taken some lessons on competitive programming, chances are that you have already heard about one of the most famous formula: the Möbius inversion. This article is aimed to provide some basic insight on what is the Möbius inversion, as well as how to apply it in various programming tasks.

Prequisite

If you are not familiar with the linear sieve and multiplicative functions, it is recommended that you read about them first here.

I will introduce some frequently used notations and lemmas first.

Notation

1. $[P]$ refers to the boolean expression, i.e. $[P] = 1$ when P is true, and 0 otherwise.
2. $\lfloor x \rfloor$ refers to rounding x down to the nearest integer. Thus $\lfloor \frac{q}{p} \rfloor$ refers to the integer division.
3. $d|n$ means that d can divide n (without a remainder).

The following functions are all multiplicative functions, where p is a prime number and k is a positive integer.

1. The constant function $I(p^k) = 1$.
2. The identity function $Id(p^k) = p^k$.
3. The power function $Id_a(p^k) = p^{ak}$, where a is constant.
4. The unit function $\epsilon(p^k) = [p^k = 1]$.
5. The divisor function $\sigma_a(p^k) = \sum_{i=0}^{k} p^{ai}$, denoting the sum of the a-th powers of all the positive divisors of the number.
6. The Möbius function $\mu(p^k) = [k = 0]$ - $[k = 1]$.
7. The Euler's totient function $\varphi(p^k) = p^k - p^{k-1}$.

Lemma

I have some my unofficial names for these frequently used conclusions. If you happen to know any more commonly used name for them, you are more than welcome to tell me.

- (**The integer division lemma**) For every positive integer p, q and r, $\lfloor \frac{\lfloor \frac{p}{q} \rfloor}{r} \rfloor = \lfloor \frac{p}{qr} \rfloor$.

Proof: We know that $\frac{p}{q} - \lfloor \frac{p}{q} \rfloor < 1$. Hence $\frac{p}{qr} - \frac{\lfloor \frac{p}{q} \rfloor}{r} < \frac{1}{r}$. Since the fraction part of $\frac{\lfloor \frac{p}{q} \rfloor}{r}$ cannot exceed $\frac{r-1}{r}$, we achieve the conclusion.

- (**The harmonic lemma**) Consider the harmonic sequence on integer division $\left(\lfloor \frac{n}{1} \rfloor, \lfloor \frac{n}{2} \rfloor, ..., \lfloor \frac{n}{n} \rfloor \right)$.

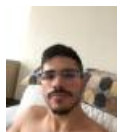1. The sequence is non-increasing, and there are at most $2\sqrt{n}$ different elements.

2. The sum of the sequence is approximate to $O(n \log n)$.

Proof: Denote $d(i) = \lfloor \frac{n}{i} \rfloor$. For every i not exceeding $\sqrt{n}$, there will be no more than $\sqrt{n}$ values in the domain of $d(i)$, so there can be at most $\sqrt{n}$ different values for $d(i)$. For the rest part of i greater than $\sqrt{n}$, we can infer that $d(i) < \sqrt{n}$ and is a positive integer, so there can be at most $\sqrt{n}$ different values for $d(i)$. Therefore we know that there are at most $2\sqrt{n}$ different elements in the sequence.

Since the Euler–Mascheroni constant states that $\lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{i} = \ln n + \gamma$, we know that $\lim_{n\to\infty} \sum_{i=1}^{n} \frac{n}{i} = n \ln n + n\gamma$, so the sum of the original sequence is approximate to $O(n \log n)$.

Moreover, it is actually possible to exploit the property that the sequence has at most $2\sqrt{n}$ different elements, and write a loop that runs in $O(\sqrt{n})$ complexity to iterate through every possible value of $d(i)$, using the fact that the greatest integer x satisfying $d(x) = d(i)$ is $\lfloor \frac{n}{\lfloor \frac{n}{i} \rfloor} \rfloor$. The piece of code below demonstrates one way to program it.

```
for (int i = 1, la; i <= n; i = la + 1) {
    la = n / (n / i);
    //n / x yields the same value for i <= x <= la.
}
```

- (**The transitive property of multiplicative functions**)

1. If both $f(x)$ and $g(x)$ are multiplicative, then $h(x) = f(x)g(x)$ is also multiplicative.
2. If both $f(x)$ and $g(x)$ are multiplicative, then their Dirichlet convolution $f * g(x) = \sum_{d|x} f(d)g(\frac{x}{d})$ is also multiplicative. Specifically, if $f(x)$ is multiplicative, $h(x) = \sum_{d|x} f(d)$ is also multiplicative.

The proof is more mathematical, which I will skip here.

What is the Möbius inversion?

According to Wikipedia, the Möbius inversion states that:

If $g(n) = \sum_{d|n} f(d)$ for every positive integer n, then $f(n) = \sum_{d|n} g(d)\mu(\frac{n}{d})$, where $\mu(x)$ is the Möbius function, which is multiplicative and satisfies $f(1) = 1, f(p) = -1, f(p^k) = 0$ for any prime number p and any integer $k \geq 2$. (*It is worth noting that you can pre-compute all values of the Möbius function from 1 to n using the linear sieve.*)

However, the definition alone does not mean much (unless the problem statement explicitly gives you something like $g(n) = \sum_{d|n} f(d)$. In that case, well...). There is one important property that is probably more useful than the definition:

$$\sum_{d|n} \mu(d) = \epsilon(n) = [n = 1]$$

In order to prove this property, we have to use the transitive property of multiplicative functions to show that $f(n) = \sum_{d|n} \mu(d)$ is multiplicative. After that, we can see $f(1) = 1$ and $f(p^k) = 0$, which means $f(x) = \epsilon(x)$. Now you may ask: why is this property important? I think it is best to show the reasons in some basic examples below.

Example 1. Find out the number of co-prime pairs of integers (x, y) in range $[1, n]$.

Solution 1. Notice that the question is the same as asking you the value of

$$f(n) = \sum_{i=1}^{n} \sum_{j=1}^{n} [gcd(i, j) = 1]$$

Now apply the Möbius inversion on $[gcd(i, j) = 1]$, we have

$$f(n) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{d|gcd(i,j)} \mu(d)$$

which is the same as

$$f(n) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{d=1}^{n}[d|gcd(i,j)]\mu(d)$$

Notice that $[d|gcd(i,j)] = [d|i][d|j]$. Therefore

$$f(n) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{d=1}^{n}[d|i][d|j]\mu(d)$$

We can change the order of summing things up, so

$$f(n) = \sum_{d=1}^{n}\mu(d)\left(\sum_{i=1}^{n}[d|i]\right)\left(\sum_{j=1}^{n}[d|j]\right)$$

We know that $\sum_{i=1}^{n}[d|i] = \sum_{j=1}^{n}[d|j] = \lfloor\frac{n}{d}\rfloor$. Thus

$$f(n) = \sum_{d=1}^{n}\mu(d)\lfloor\frac{n}{d}\rfloor^2$$

Then we can simply loop through 1 to n to compute the formula. While we can optimize this loop to $O(\sqrt{n})$ using the harmonic lemma, we still have to use the linear sieve to pre-compute the values of $\mu(d)$. Therefore, the overall complexity of the algorithm is $O(n)$.

Example 2. Find out the sum of $gcd(x, y)$ for every pair of integers (x, y) in range $[1, n]$ ($gcd(x, y)$ means the greatest common divisor of x and y).

Solution 2. Similar as above, notice that the question is the same as asking you the value of

$$f(n) = \sum_{i=1}^{n}\sum_{j=1}^{n} gcd(i,j)$$

Let $k = gcd(i, j)$. We can then loop for k first.

$$f(n) = \sum_{k=1}^{n}k\sum_{i=1}^{n}\sum_{j=1}^{n}[gcd(i,j) = k]$$

Now assume $i = ak, j = bk$, and then

$$f(n) = \sum_{k=1}^{n}k\sum_{a=1}^{\lfloor\frac{n}{k}\rfloor}\sum_{b=1}^{\lfloor\frac{n}{k}\rfloor}[gcd(a,b) = 1]$$

It can be observed that the last part of the formula is the same as the one in example 1. Applying the same procedure, we have

$$f(n) = \sum_{k=1}^{n}k\sum_{d=1}^{\lfloor\frac{n}{k}\rfloor}\mu(d)\lfloor\frac{n}{kd}\rfloor^2$$

According to the harmonic lemma, $O(\sum_{k=1}^{n}\lfloor\frac{n}{k}\rfloor) = O(n\log n)$. Therefore, if we compute the sum above with brute force, the overall complexity will be $O(n\log n)$.

This is, however, not the best complexity we can achieve with the Möbius inversion. Now let $l = kd$, and we can loop for l first.

$$f(n) = \sum_{l=1}^{n}\lfloor\frac{n}{l}\rfloor^2\sum_{k|l}k\mu(\frac{l}{k})$$

Applying the transitive property of multiplicative functions, we can see that $g(l) = \sum_{k|l}k\mu(\frac{l}{k})$ is multiplicative. Moreover, we have $g(p^k) = p^k - p^{k-1}$. If you have

somewhat studied about the number theory, you may figure out that $g(l)$ is actually the Euler's totient function. Regardless of whether you know about the coincidence or not, $g(l)$ can be pre-computed with the linear sieve, and $f(n) = \sum_{l=1}^{n} \lfloor \frac{n}{l} \rfloor^2 g(l)$, which can be simply computed in $O(n)$ complexity.

Example 3. Find out the sum of $lcm(x, y)$ for every pair of integers (x, y) in range $[1, n]$ ($lcm(x, y)$ means the least common multiple of x and y).

Solution 3. Well, we should be pretty familiar with the techniques by now. First of all, the answer should be

$$f(n) = \sum_{i=1}^{n} \sum_{j=1}^{n} lcm(i, j)$$

Since $lcm(i, j) = \frac{ij}{gcd(i,j)}$, let $k = gcd(i, j)$. We can then loop for k first.

$$f(n) = \sum_{k=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} [gcd(i, j) = k] \frac{ij}{k}$$

Now assume $i = ak, j = bk$, and then

$$f(n) = \sum_{k=1}^{n} \sum_{a=1}^{\lfloor \frac{n}{k} \rfloor} \sum_{b=1}^{\lfloor \frac{n}{k} \rfloor} [gcd(a, b) = 1]abk$$

$$= \sum_{k=1}^{n} k \sum_{a=1}^{\lfloor \frac{n}{k} \rfloor} a \sum_{b=1}^{\lfloor \frac{n}{k} \rfloor} b[gcd(a, b) = 1]$$

$$= \sum_{k=1}^{n} k \sum_{a=1}^{\lfloor \frac{n}{k} \rfloor} a \sum_{b=1}^{\lfloor \frac{n}{k} \rfloor} b \sum_{d=1}^{\lfloor \frac{n}{k} \rfloor} [d|a][d|b]\mu(d)$$

$$= \sum_{k=1}^{n} k \sum_{d=1}^{\lfloor \frac{n}{k} \rfloor} \mu(d) \left(\sum_{a=1}^{\lfloor \frac{n}{k} \rfloor} [d|a]a \right) \left(\sum_{b=1}^{\lfloor \frac{n}{k} \rfloor} [d|b]b \right)$$

Now let's proceed to assume $a = dp, b = dq$, then

$$f(n) = \sum_{k=1}^{n} k \sum_{d=1}^{\lfloor \frac{n}{k} \rfloor} \mu(d) \left(\sum_{p=1}^{\lfloor \frac{n}{kd} \rfloor} dp \right) \left(\sum_{q=1}^{\lfloor \frac{n}{kd} \rfloor} dq \right)$$

$$= \sum_{k=1}^{n} k \sum_{d=1}^{\lfloor \frac{n}{k} \rfloor} \mu(d)d^2 \left(\sum_{p=1}^{\lfloor \frac{n}{kd} \rfloor} p \right) \left(\sum_{q=1}^{\lfloor \frac{n}{kd} \rfloor} q \right)$$

Now notice $\sum_{p=1}^{\lfloor \frac{n}{kd} \rfloor} p = \frac{(1+\lfloor \frac{n}{kd} \rfloor)(\lfloor \frac{n}{kd} \rfloor)}{2}$, and we have

$$f(n) = \sum_{k=1}^{n} k \sum_{d=1}^{\lfloor \frac{n}{k} \rfloor} \mu(d)d^2 \left(\frac{(1 + \lfloor \frac{n}{kd} \rfloor)(\lfloor \frac{n}{kd} \rfloor)}{2} \right)^2$$

Now following the example above, let $l = kd$, and then

$$f(n) = \sum_{l=1}^{n} \left(\frac{(1 + \lfloor \frac{n}{l} \rfloor)(\lfloor \frac{n}{l} \rfloor)}{2} \right)^2 \sum_{d|l} \mu(d)ld$$

Notice that $g(l) = \sum_{d|l} \mu(d)ld$ is also multiplicative, and $g(p^k) = p^k - p^{k+1}$. We can pre-compute the values of $g(l)$ via the linear sieve, and $f(n) = \sum_{l=1}^{n} \left(\frac{(1+\lfloor \frac{n}{l} \rfloor)(\lfloor \frac{n}{l} \rfloor)}{2} \right)^2 g(l)$. The overall complexity will be $O(n)$.

Example 4. Find out the sum of $lcm(A[i], A[j])$ for every pair of integers $(A[i], A[j])$ with an array A of length N ($1 \leq i, j \leq N$), with the constraint of $1 \leq N \leq 200,000$ and $1 \leq A[i] \leq 1,000,000$.

Solution 4. This is a common variation of the example above. Denote

$$f(n) = \sum_d \sum_{a \in A \wedge d|a} \sum_{b \in A \wedge d|b} [gcd(a,b) = d] \frac{ab}{d}$$

We simply have

$$f(n) = \sum_d \sum_{a \in A \wedge d|a} \sum_{b \in A \wedge d|b} [gcd(\frac{a}{d}, \frac{b}{d}) = 1] \frac{ab}{d}$$

$$= \sum_d \sum_{a \in A \wedge d|a} \sum_{b \in A \wedge d|b} \sum_l [l|\frac{a}{d}][l|\frac{b}{d}] \mu(l) \frac{ab}{d}$$

$$= (\sum_d \frac{1}{d} \sum_l \mu(l))(\sum_{a \in A} \sum_{b \in A} [dl|a][dl|b] ab)$$

Let $t = dl$, then

$$f(n) = (\sum_t \sum_{l|t} \frac{l}{t} \mu(l))(\sum_{a \in A \wedge t|a} a)^2$$

Now, similar to the example above $g(t) = \sum_{l|t} \frac{l}{t} \mu(l)$ is multiplicative with

$g(p^k) = p^{-k} - p^{-k+1}$, and thus can be computed with $O(V)$, where V is the range span of

A. The latter part of the formula $\left(\sum_{a \in A \wedge t|a} a \right)^2$ is also fairly easy to figure out if you keep track

of the number of the elements equal to v as $cnt(v)$, since

$$\sum_{a \in A \wedge t|a} a = \sum_{t|v} cnt(v)v$$

$$= \sum_{i=1}^{\frac{V}{t}} cnt(it)it$$

and with the harmonic lemma, the above computation can be done in $O(V \log V)$. Hence
the overall complexity is $O(N + V \log V)$.

Practice Problems

Simple Sum

(Thanks **Bashca** for showing me this one!)

Link

GCD

Link

Sky Code

Link

Extensions

It is known that all examples above can be further optimized to $O(n^{\frac{2}{3}})$. The exact technique
will (probably) be introduced in the upcoming article.

Update: The article introducing the optimization is ready here.