CDM

Möbius Inversion

Klaus Sutner
Carnegie Mellon University

24-moebius 2017/12/15 23:16

1 Möbius Inversion

- Algebraic Proof

Necklaces

We already know how to deal with counting problems relating to necklaces and bracelets.

Definition

A k-ary necklace is a word over a k symbol alphabet up to rotation. A k-ary bracelet is a word over a k symbol alphabet up to rotation and reflection.

If you don't like words, think colored beads.

If the words have length n then the appropriate actions are given by the cyclic group $\mathbb{Z}_n$ and the dihedral group D_n, respectively.

No problem.

Primitive Words

But here is a very closely related notion that produces a new type of counting problem.

Definition

A non-empty word w is primitive if it is not of the form x^i for any x shorter than w. The root of a word w is the shortest word x such that $x^i = w$ and the exponent i is then the repetition factor of w.

Thus a primitive word is its own root and has repetition factor 1.

A word of prime length is primitive unless it is of the form a^n.

Primitive words pop up naturally when one tries to understand commutativity in words.

A lemma

Lemma

Suppose u and v are non-empty words such that $uv = vu$.

Then u and v have the same root.

Proof.

Use induction on $|uv|$, the combined length of u and v.

And draw a picture.

□

In other words, words never commute except in the trivial case.

Counting Primitive Words

Even if n has many factors, random words of length n tend to be primitive.

In order to count primitive words, let $\pi(n)$ denote the number of primitive words of length n over an alphabet of size k. Note that

$$k^n = \sum_{d|n} \pi(d)$$

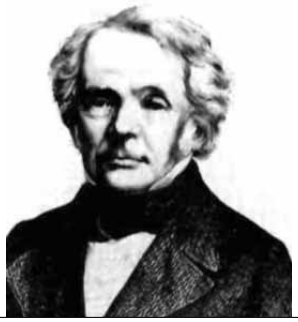
Wishful thinking: It would be nice if we could turn this equation around:

We would like to obtain an expression for $\pi(n)$ in terms of k^n.

Möbius Function

There is a well-known method to tackle problems of this form based on the Möbius function μ which is defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^r & \text{if } n = p_1 \dots p_r, \text{ all distinct primes,} \\ 0 & \text{if } n \text{ is divisible by } p^2, \ p \text{ a prime.} \end{cases}$$



Möbius Inversion

Lemma (Möbius Inversion Formula)

Let $f(n) = \sum_{d|n} g(d)$. *Then*

$$g(n) = \sum_{d|n} \mu(d) f(n/d) = \sum_{d|n} \mu(n/d) f(d).$$

In other words, if f is obtained from g by summation (over divisors), then g can in turn be expressed by summation over f.

Why On Earth???

Proposition

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Proof.

Suppose $n > 1$ with prime decomposition

$$n = p_1^{e_1} p_2^{e_2} \dots p_k^{e_k}$$

Let $n' = p_1 p_2 \dots p_k$. Clearly $\sum_{d|n} \mu(d) = \sum_{d|n'} \mu(d)$. But the latter sum is

$$\sum_{I \subseteq [k]} (-1)^{|I|} = \sum_{\ell \le k} (-1)^\ell \binom{k}{\ell} = 0.$$

Done. □

Calculating

$$\begin{aligned} \sum_{d|n} \mu(d) f(n/d) &= \sum_{d|n} \mu(n/d) f(d) \\ &= \sum_{d|n} \mu(n/d) \sum_{e|d} g(e) \\ &= \sum_{e|d|n} \mu(n/d)\, g(e) \\ &= \sum_{e|n} \sum_{d|n/e} \mu(n/ed)\, g(e) \\ &= \sum_{e|n} \sum_{d|n/e} \mu(d)\, g(e) \\ &= g(n) \end{aligned}$$

Standard Application: Euler's Totient Function

Recall Euler's totient function

$$\varphi(n) = \#(\, k \mid 1 \le k < n, \gcd(k, n) = 1)$$

Then

$$\sum_{d|n} \varphi(d) = n$$

Hence

$$\varphi(n) = \sum_{d|n} \mu(d) n/d$$

No Free Lunch

Alas, computationally this is not a great step forward:

- To compute φ directly, we need prime factorization.
- To compute φ via μ, we need prime factorization.

This is inevitable in a way: since factorization is presumably hard, and division is cheap, the complexity has to go somewhere else: it hides in μ.

Computationally there is no free lunch, ever. Maybe occasionally breakfast.

Application: Primitive Words

Fix an alphabet of size k. For the number of primitive words the Möbius inversion formula yields

$$\pi(n) = \sum_{d|n} \mu(d) k^{n/d}$$

Over a binary alphabet the number of primitive words up to length 12 is

$$2, 2, 6, 12, 30, 54, 126, 240, 504, 990, 2046, 4020$$

Application: Necklaces

Define $\text{neck}(n, k)$ to be the number of necklaces of length n using beads of k colors. By stringing up primitive words into necklaces we get

$$\begin{aligned}
\text{neck}(n,k) &= \sum_{d|n} \frac{1}{d} \pi(d) \\
&= \frac{1}{n} \sum_{d|n} \frac{n}{d} \sum_{e|d} \mu(e) k^{d/e} \\
&= \frac{1}{n} \sum_{d|n} \sum_{e|n/d} \mu(e) d\, k^{n/de} \\
&= \frac{1}{n} \sum_{D|n} \sum_{ed=D} \mu(e) D/e\, k^{n/D} \\
&= \frac{1}{n} \sum_{D|n} \varphi(D) k^{n/D}
\end{aligned}$$

Lyndon Words

In each necklace (considered as an equivalence class of words) we can pick the lexicographically first as a representative.

Definition

A Lyndon word is a representative for a necklace that is also primitive.

Proposition

The number of Lyndon words of length n over a k symbol alphabet is

$$\frac{1}{n} \pi(n) = \frac{1}{n} \sum_{d|n} \mu(n/d) k^d.$$

- Möbius Inversion

2. Algebraic Proof

Möbius Meets Algebra

What is really going on? Our proof, while correct, offers no insights. A good proof should explain why the result holds, not just establish its formal correctness.

Definition

An arithmetic function is a function $f : \mathbb{N}^+ \to \mathbb{C}$ from the positive integers to the complex numbers.

The Dirichlet product or convolution of two arithmetic functions f and g is defined as

$$(f * g)(n) = \sum_{d|n} f(d) g(n/d) = \sum_{xy=n} f(x) g(y).$$

Thus, an arithmetic function is just an infinite sequence of numbers. Often the numbers are integers or rationals, but we might as will deal with the more general case.

Properties of the Dirichlet Product

Let ε by the arithmetic function defined by

$$\varepsilon(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

The following is easy to see.

Proposition

The Dirichlet product is associative: $f * (g * h) = (f * g) * h$ and commutative: $f * g = g * f$.

Moreover, ε is a neutral element: $f * \varepsilon = \varepsilon * f = f$.

So, we are dealing with a commutative monoid.

Inverse Elements

Which arithmetic functions f have an inverse f^{-1} such that For example $\varepsilon^{-1} = \varepsilon$.

Lemma

f has an inverse iff $f(1) \neq 0$.

Proof.

Suppose g is the inverse of f. Then

$$1 = (f * g)(1) = f(1)g(1)$$

so that $f(1) \neq 0$.

For the opposite direction let $f(1) \neq 0$. We can construct the inverse g directly by induction.

$$g(n) = \begin{cases} \frac{1}{f(1)} & \text{if } n = 1, \\ \frac{-1}{f(1)} \sum_{1<d|n} f(d)g(n/d) & \text{otherwise.} \end{cases}$$

□

Zeta and Möbius Function

Definition

The zeta function is defined by $\zeta(n) = 1$.

Its inverse is called the Möbius function μ.

This is justified by the lemma; ζ is indeed invertible.

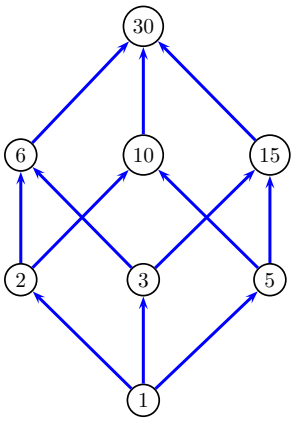
Note that convolution with ζ is essentially just a summation

$$(f * \zeta)(n) = \sum_{d|n} f(d)$$

so with $g = f * \zeta$ we have $f = g * \mu$: exactly what we wanted.

Partial Orders

Since we are not summing over intervals $i = 1, \ldots, n$ but over the divisors of n we need to talk about partial orders.



Intervals

For any partial order P define intervals by

$$[a, b] = \{ x \in P \mid a \leq x \leq b \}$$

and int(P) the collection of all non-empty intervals over P.

Now consider two functions $f, g : \text{int}(P) \to \mathbb{R}$. The convolution of f and g is

$$(f * g)([a, b]) = \sum_{a \leq x \leq b} f([a, x])g([x, b])$$

Technicalities

For this to work we have to assume that P is locally finite: all intervals $[a, b]$ must be finite.

Also note that these intervals are not necessarily linearly ordered, the picture from above is the interval $[1, 30]$.

The notation $f([a, b])$ is technically correct, but really an atrocity.

One usually just writes $f(a, b)$ instead.

Incidence Algebra

Definition

$\langle \text{int}(P), * \rangle$ is the incidence algebra (over the partial order P).

The delta function is defined to be $\delta(x, x) = 1$, and 0 otherwise.

Proposition

The incidence algebra is a monoid: convolution is an associative operation, and δ is a neutral element.

Proof is a standard exercise in summation.

Inverse Elements

One might wonder what elements possess an inverse in the incidence algebra:

$$f * g = g * f = \delta$$

and what these inverses might look like.

Lemma

An element f of the incidence algebra is invertible iff $f(x,x) \neq 0$ for all x.

Proof. Necessity is easy.

$$1 = \delta(x,x) = (f*g)(x,x) = f(x,x)g(x,x).$$

Proof, contd.

For sufficiency we will define g explicitly.

$$g(x,y) = \begin{cases} \frac{1}{f(x,x)} & \text{if } x = y, \\ \frac{-1}{f(x,x)} \sum_{x<z\leq y} f(x,z)g(z,y) & \text{otherwise.} \end{cases}$$

This definition uses induction on the size of the intervals. By brute force, $f * g = \delta$.

Note that g also has the property from the lemma, so let h be its right inverse.

$$g * f = (g * f) * (g * h) = g * (f * g) * h = g * h = \delta.$$

So, g is also a left inverse.

□

Zeta Function

Definition

The zeta function is defined by $\zeta(x,y) = 1$.

From the lemma, ζ is invertible.

Its inverse is called the Möbius function μ of int(P).

Note that convolution with ζ is essentially just a summation

$$f * \zeta(a,b) = \sum_{a\leq z\leq b} f(a,z)$$

so with $g = f * \zeta$ we have $f = g * \mu$: exactly what we wanted.

The theorem

Theorem

Let $g(x,y) = \sum_{x\leq z\leq y} f(x,z)$. Then

$$f(x,y) = \sum_{x\leq z\leq y} \mu(z,y) \cdot g(x,z).$$

This is very pretty, but the problem is that we need to compute μ.

There are techniques to do this for interesting cases (using products of partial orders), but we won't go there.

For the divisibility order on the natural numbers we get the classical Möbius function – sort of.

Another Technicality

In our applications, the partial order is divisibility on the natural numbers.