MÖBIUS INVERSION FORMULA

ERIN STUHLSATZ

1. Introduction

Many problems in mathematics, specifically in combinatorics, can be simplified by a simple change of perspective. Often a difficult counting problem can be solved when the sets are counted in different ways [3]. One tool for proving difficult theorems is the Möbius Inversion Formula, as shown below.

To begin, we'll review some important definitions and theorems from Set Theory. These are used in the definition and proof of the Principle of Inclusion and Exclusion (or PIE). The PIE is fairly commonly used, most obviously in problems such as that in Section 3.1.1, where, given the number of people in a room with certain characteristics, we use the PIE to compute the number of people who don't share those characteristics.

The PIE can be proved by induction using oly the Set Theory explained in Section 2. Alternatively, the Möbius Inversion Formula gives a more elegant proof. In sections 5 through 8, we'll focus on the Möbius Inversion Formula and Möbius Functions in general.

Möbius Functions are piecewise functions defined on any set, partially ordered by some relation. We'll derive Möbius Functions for a set consisting of sets ordered by inclusion, a set of arborescences ordered, and the set of integers ordered by division. We'll then examine the Upside-Down Möbius Function, where the relation is defined backwards; that is, if $x \leq y$, we would say $y \leq^* x$. Finally, we'll prove the Möbius Inversion Formula by induction.

The Möbius Inversion Formula states that given any function $f(x)$, if we define $g(x)$ as follows:

$$g(x) = \sum_{z \leq x} f(z),$$

then

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

Date: 4 April.

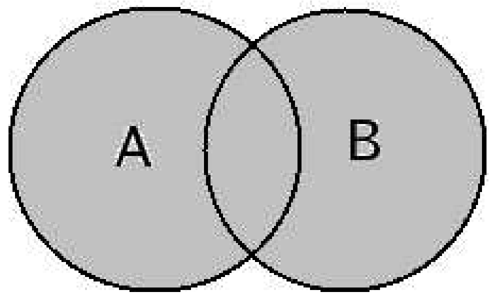

Figure 1. A "union" B, $A \cup B$

By appropriately defining $f(x)$, we can use the Möbius Inversion Formula to prove the Principle of Inclusion and Exclusion.

2. Very Basic Review

2.1. **Definitions.** A **set** is a collection of objects. These objects can be anything from numbers to letters to chickens. The number of objects within a set is the **size** of the set. If some object x is a member of the set A, we write $x \in A$. The size of the set A is denoted $|A|$. Some sets, such as the set of whole numbers, are infinite; however, the sets used in this paper are finite. The set with size 0, or the **empty set**, is denoted $\emptyset$.

Set Theory requires its own set of operations to define how sets interact.

In Figures 1-5, we denote the elements in sets A and B by the points inside the circles labeled A and B respectively.

The set consisting of all the elements that are members of A, B, or both is called the **union** of the two sets A and B, or "A union B", and is denoted $A \cup B$. In Figure 1, $A \cup B$ consists of the shaded region. This operation can be generalized to more sets. The union of n sets A_1, A_2, ..., A_n, denoted $A_1 \cup A_2 \cup \cdots \cup A_n$ is the set consisting of all elements x where $x \in A_i$ for at least one $i \leq n$.

The set of all the elements that are members of both A and B is called the **intersection** of A and B, or "A intersect B" and is denoted $A \cap B$. The shaded region in Figure 2 is the intersection between the sets. Intersection can also be generalized to more sets. The intersection of n sets A_1, A_2, ..., A_n, denoted $A_1 \cap A_2 \cap \cdots \cap A_n$ is the set consisting of all elements that are in every A_i.

We say a set A is a **subset** of another set B, or $A \subseteq B$, if every element in A is also a member of B. In Figure 3 the set A is a subset of the set B, since A is completely contained within B.

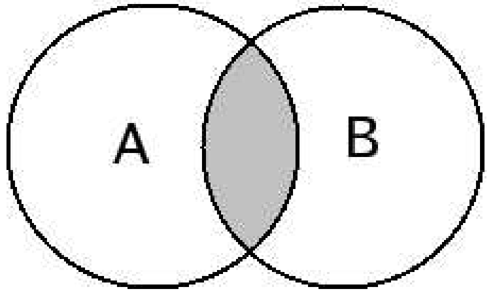

FIGURE 2. A "intersect" B, $A \cap B$

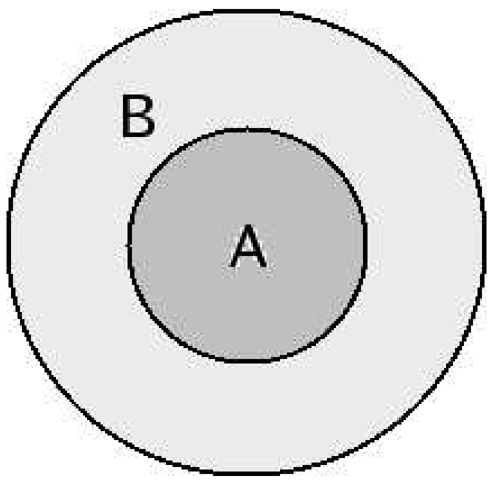

FIGURE 3. A is a subset of B, $A \subseteq B$

Two sets A and B are **equal** ($A = B$) if they have all the same elements. This implies that every element of A is also an element of B, and every element of B is also an element of A; that is, both sets are subsets of each other.

Where A is a subset of some other set S, the **complement** of set A, denoted A', is the set of all points in S that are not in A. A' is shown as the shaded section in Figure 4. The set of elements that are in A but not in B, shown as the shaded section in Figure 5, is the **difference** of the two sets A and B, denoted $A - B$. $A - B$ is also the intersection of the sets A and B', denoted $A \cap B'$.

Consider A and B to be subsets of the set S={1,2,3,...,10} with A={1,2,3,4} and B={2,4,6,8}.

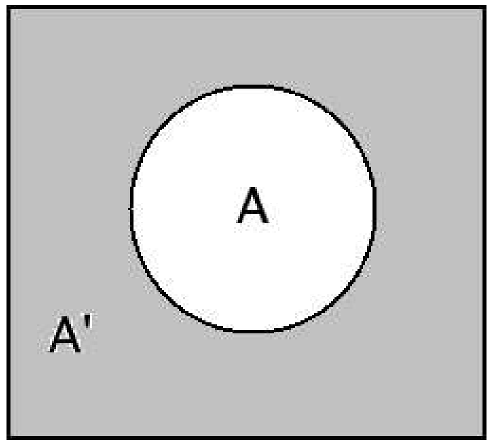

FIGURE 4. "Compliment" of A, A'

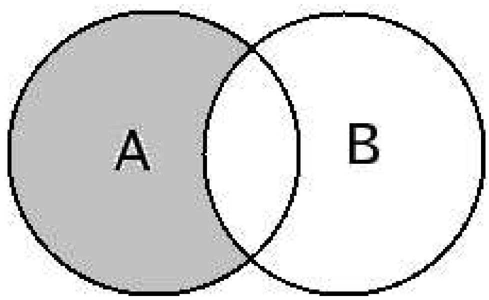

FIGURE 5. $A - B$, $A \cap B'$

Then $A \subseteq S$,

$$\begin{aligned} A \cup B &= \{1,2,3,4,6,8\}, \\ A \cap B &= \{2,4\}, \\ A' &= \{5,6,7,8,9,10\}, \\ S - A &= \{5,6,7,8,9,10\}, \text{ and} \\ A' &= S - A. \end{aligned}$$

2.2. **Laws.** When dealing with operations on the real numbers, there are some basic laws that we take for granted, such as the Associative Law, which states that for any real numbers x, y, and z, we have $(x + y) + z = x + (y + z)$. Two important laws for operations on sets are the Associative and Distributive Laws.

Associative Law for Union. *The order in which unions are calculated doesn't affect the final result.*

$$(A \cup B) \cup C = A \cup (B \cup C).$$

Proof

In order to prove that two sets are equal, we need to show that each set is a subset of the other. If some element x is a member of the first set, it must also be a member of the second, and vice versa.

First, assume that

$$x \in (A \cup B) \cup C.$$

From the definition of the union, either $x \in A \cup B$ or $x \in C$. If the former is true, then either $x \in A$ or $x \in B$. This means x is a member of at least one of the sets A, B, or C.

Suppose $x \in B$ or $x \in C$. Then $x \in B \cup C$, since this set consists of every element that is a member of B, C, or both. Since $x \in B \cup C$, then $x \in A \cup (B \cup C)$.

Suppose $x \in A$, then $x \in A \cup (B \cup C)$.

Thus, $x \in (A \cup B) \cup C$ implies $x \in A \cup (B \cup C)$, and we have shown

$$(A \cup B) \cup C \subseteq A \cup (B \cup C).$$

Now, assume $y \in A \cup (B \cup C)$. Thus, either $y \in A$ or $y \in B \cup C$.

Suppose $y \in B \cup C$. Then either $y \in B$ or $y \in C$. So once again, we have that y is a member of at least one of the sets A, B, or C.

If y is in A or B, then it will be a member of $A \cup B$, and therefore a member of $(A \cup B) \cup C$. If $y \in C$, then we know $y \in (A \cup B) \cup C$.

Thus, $y \in A \cup (B \cup C)$ implies $y \in (A \cup B) \cup C$, and

$$A \cup (B \cup C) \subseteq (A \cup B) \cup C.$$

Since the two sets are subsets of each other, they must be equal. Therefore,

$$A \cup (B \cup C) = (A \cup B) \cup C.$$

Since the order in which we compute unions doesn't affect the final result, the notation $A_1 \cup A_2 \cup \cdots \cup A_n$ is not ambiguous.

Associative Law for Intersection. *The order in which intersections are computed doesn't affect the final set; that is,*

$$(A \cap B) \cap C = A \cap (B \cap C).$$

Proof

We'll proceed by the same method for this law.

First, assume

$$x \in (A \cap B) \cap C.$$

Since this is an intersection, this means x is a member of both $A \cap B$ and C. Since x is a member of $A \cap B$, it must be a member of both A and B as well. So the following are all true:

$$x \in A,$$

$$x \in B, \qquad \text{and}$$
$$x \in C.$$

Since x is a member of both B and C, we have $x \in B \cap C$. Since it's a member of A as well, $x \in A \cap (B \cap C)$. Therefore, any x that is a member of the first set is also a member of the second set, and

$$(A \cap B) \cap C \subseteq A \cap (B \cap C).$$

Now suppose that y is a member of $A \cap (B \cap C)$. Thus, y is a member of both A and $B \cap C$. Since y is a member of $B \cap C$, it must be a member of both B and C. So, we have that y is a member of all three sets A, B, and C. Since y is a member of both A and B, $y \in A \cap B$. Since $y \in C$ as well, $y \in (A \cap B) \cap C$. Therefore, any y that is a member of the second set is also a member of the first set, and

$$A \cap (B \cap C) \subseteq (A \cap B) \cap C.$$

The two sets are subsets of each other, and

$$(A \cap B) \cap C = A \cap (B \cap C).$$

The Distributive Law for algebra states that $x(y + z) = xy + xz$. The Distributive Law for sets uses "intersection" and "union" instead of multiplication and addition, but the result is similar.

Distributive Law. *Given any sets A, B, and C,*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Proof

First, suppose that

$$x \in A \cap (B \cup C).$$

So, $x \in A$ and $x \in B \cup C$. Since $x \in B \cup C$, either $x \in B$ or $x \in C$. If $x \in B$, then x is a member of both sets A and B, and $x \in A \cap B$. If x is a member of C, then it is a member of $A \cap C$. So, either $x \in A \cap B$ or $x \in A \cap C$. Thus, $x \in (A \cap B) \cup (A \cap C)$. Therefore,

$$A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C).$$

Now, suppose that

$$y \in (A \cap B) \cup (A \cap C).$$

Either $y \in A \cap B$ or $y \in A \cap C$. If $y \in A \cap B$, then y is in A and y is in B. If $y \in A \cap C$, then $y \in A$ and $y \in C$. So y is definitely in the set A, and either $y \in B$ or $y \in C$. Since y is in either B or C, we have $y \in B \cup C$. Since it's in both A and $B \cup C$, thus $y \in A \cap (B \cup C)$. Therefore,

$$(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C).$$

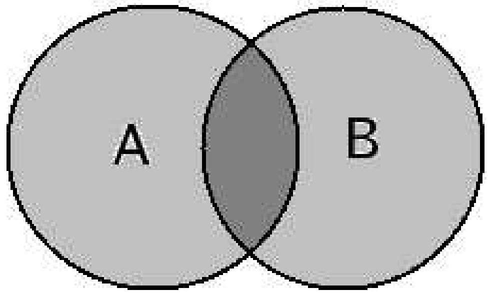

FIGURE 6. $|A \cup B|$

Since these two sets are subsets of each other, they must be equal. Therefore,

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

3. Principle of Inclusion and Exclusion

The Principle of Inclusion and Exclusion (PIE) is used to calculate the size of the union of finite sets. We will use the notation PIE_n to denote the Principle of Inclusion and Exclusion on n sets.

3.1. **PIE_2: The Principle of Inclusion and Exclusion for two sets.** In a very simple example, we examine the union of two sets A and B. We can't calculate the size of the union by simply adding the sizes of the two sets, since they may have elements in common. In Figure 6, if we were to add the sizes of the two sets, the lightly shaded sections would each be counted once. However, the darker section would be counted twice, once as a part of each set. Thus, to ensure that the repeated terms are only counted once, the PIE_2 states that given finite sets A and B,

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

The following example, which demonstrates an application of the PIE_2, comes from *Introduction to Discrete Mathematics* [1].

3.1.1. *Example.* There are 100 people in a room. In this group, 60 are men, 30 are young, and 10 are young men. How many are old women?

Let P be the set of people in the room: $|P| = 100$. Let M be the set of all men in the room, $|M| = 60$, and let Y be the set of all young people in the room,$|Y| = 30$. Then the set of young men will be the set $M \cap Y$, and $|M \cap Y| = 10$. We want to know the size of $M' \cap Y'$, or the set of all people who are both not young and not men. Thus, we could either count the set of all old women, or we could count the number of people who are either men or young, and subtract that from the total number of people. This is the same as calculating $|P| - |M \cup Y|$.

By the Principle of Inclusion and Exclusion,

$$|M \cup Y| = |M| + |Y| - |M \cap Y|,$$

So,

$$\begin{aligned} |S| - |M \cup Y| &= |S| - |M| - |Y| + |M \cap Y| \\ &= 100 - 60 - 30 + 10 \\ &= 20 \end{aligned}$$

There are 20 old women in the room.

3.2. **PIE_3.** The proof for the Principle of Inclusion and Exclusion with three sets introduces the principles used to prove the general case of the PIE (PIE_n).

Principle of Inclusion-Exclusion for 3 sets. *Given finite sets A, B, and C*

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

Proof Suppose we want to calculate the size of the union of three sets A, B, and C. By the Associative Law, we can calculate the union of three sets in any order we wish, so $A \cup B \cup C = A \cup (B \cup C)$. Thus,

$$|A \cup B \cup C| = |A \cup (B \cup C)|.$$

The PIE_2 gives the size of the union of any two sets, so

$$|A \cup (B \cup C)| = |A| + |B \cup C| - |A \cap (B \cup C)|$$

We can use the PIE_2 to find the size of $B \cup C$, and we can use the Distributive Law to expand $A \cap (B \cup C)$.

$$|A \cup (B \cup C)| = |A| + |B| + |C| - |B \cap C| - |(A \cap B) \cup (A \cap C)|.$$

This last term is another instance of a union between two sets, so we can use the PIE_2 once again.

$$|(A \cap B) \cup (A \cap C)| = |A \cap B| + |A \cap C| - |(A \cap B) \cap (A \cap C)|.$$

By the Distributive Law, the very last term is equal to $A \cap B \cap C$. Therefore, we can reorganize to find

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

3.3. **PIE$_n$.** We can prove the generalized Principle of Inclusion and Exclusion using induction. We've already proved the base case, for PIE$_2$.

To begin, assume that the PIE holds true for n finite sets.

Principle of Inclusion-Exclusion for n sets. *Given any n sets B_1, B_2, ..., B_n, we assume by induction that*

$$\begin{aligned}|B_1 \cup B_2 \cup \cdots \cup B_n| = \ & |B_1| + |B_2| + \cdots |B_n| \\ & -(|B_1 \cap B_2| + |B_1 \cap B_3| + \cdots + |B_{n-1} \cap B_n|) \\ & + |B_1 \cap B_2 \cap B_3| + \cdots + |B_{n-2} \cap B_{n-1} \cap B_n| \\ & \vdots \\ & +(-1)^{n-1}|B_1 \cap B_2 \cap \cdots \cap B_n|.\end{aligned}$$

We want to show that the PIE will also be true for $n+1$ finite sets; that is, given A_1, A_2,..., A_{n+1},

$$\begin{aligned}|A_1 \cup \cdots \cup A_n \cup A_{n+1}| = \ & |A_1| + |A_2| + \cdots |A_n| + |A_{n+1}| \\ & -(|A_1 \cap A_2| + |A_1 \cap A_3| + \cdots + |A_n \cap A_{n+1}|) \\ & + |A_1 \cap A_2 \cap A_3| + \cdots |A_{n-1} \cap A_n \cap A_{n+1}| \\ & \vdots \\ & +(-1)^n|A_1 \cap A_2 \cap \cdots \cap A_n \cap A_{n+1}|.\end{aligned}$$

By the Associative Law,

$$A_1 \cup A_2 \cup \cdots \cup A_{n+1} = (A_1 \cup A_2 \cup \cdots \cup A_n) \cup A_{n+1}.$$

If we consider $A_1 \cup A_2 \cup \cdots \cup A_n$ to be one set and A_{n+1} another, then

$$|(A_1 \cup A_2 \cup \cdots A_n) \cup A_{n+1}| = |A_1 \cup A_2 \cup \cdots \cup A_n| + |A_{n+1}| - |(A_1 \cup A_2 \cup \cdots \cup A_n) \cap A_{n+1}|.$$

To expand the last term, we can use the Distributive Law repeatedly.

$$(A_1 \cup A_2 \cup \cdots \cup A_n) \cap A_{n+1} = (A_1 \cap A_{n+1}) \cup (A_2 \cap A_{n+1}) \cup \cdots \cup (A_n \cap A_{n+1})$$

We define B_i to be the intersection of A_i and A_{n+1};

$$\begin{aligned}B_1 &= A_1 \cap A_{n+1} \\ B_2 &= A_2 \cap A_{n+1} \\ &\vdots \\ B_n &= A_n \cap A_{n+1}.\end{aligned}$$

Therefore,

$$(A_1 \cup A_2 \cup \cdots \cup A_n) \cap A_{n+1} = B_1 \cup B_2 \cup \cdots \cup B_n.$$

Since the union of all the sets B_i is the union of n sets, we have assumed

$$\begin{aligned}|B_1 \cup B_2 \cup \cdots \cup B_n| = \ & |B_1| + |B_2| + \cdots |B_n| \\ & -(|B_1 \cap B_2| + |B_1 \cap B_3| + \cdots + |B_{n-1} \cap B_n|) \\ & + |B_1 \cap B_2 \cap B_3| + |B_1 \cap B_2 \cap B_4| + \cdots |B_{n-2} \cap B_{n-1} \cap B_n| \\ & \vdots \\ & +(-1)^{n-1}|B_1 \cap B_2 \cap \cdots \cap B_n|.\end{aligned}$$

Thus,

$$\begin{aligned}|A_1 \cup A_2 \cup \cdots \cup A_{n+1}| = \ & |A_1| + |A_2| + \cdots |A_n| + |A_{n+1}| \\ & -(|A_1 \cap A_2| + |A_1 \cap A_3| + \cdots + |A_{n-1} \cap A_n|) \\ & + |A_1 \cap A_2 \cap A_3| + \cdots |A_{n-2} \cap A_{n-1} \cap A_n| \\ & \vdots \\ & +(-1)^{n-1}|A_1 \cap A_2 \cap \cdots \cap A_n| \\ & + |B_1| + |B_2| + \cdots |B_n| \\ & -(|B_1 \cap B_2| + |B_1 \cap B_3| + \cdots + |B_{n-1} \cap B_n|) \\ & + |B_1 \cap B_2 \cap B_3| + \cdots |B_{n-2} \cap B_{n-1} \cap B_n| \\ & \vdots \\ & +(-1)^{n-1}|B_1 \cap B_2 \cap \cdots \cap B_n|.\end{aligned}$$

If we substitute $A_i \cap A_{n+1}$ back in for B_i and reorganize, we're left with

$$\begin{aligned}|A_1 \cup A_2 \cup \cdots \cup A_n \cup A_{n+1}| = \ & |A_1| + |A_2| + \cdots |A_n| + |A_{n+1}| \\ & -(|A_1 \cap A_2| + |A_1 \cap A_3| + \cdots + |A_n \cap A_{n+1}|) \\ & + |A_1 \cap A_2 \cap A_3| + \cdots + |A_{n-1} \cap A_n \cap A_{n+1}| \\ & \vdots \\ & +(-1)^{n}|A_1 \cap A_2 \cap \cdots \cap A_n \cap A_{n+1}|.\end{aligned}$$

Since the PIE_{n+1} holds true when we assume PIE_n, we've proven the generalized Principle of Inclusion and Exclusion for n sets by induction.

The following example from Lovász shows how the PIE can be used to solve more complicated problems [2].

3.3.1. *Example.* Find a formula for $\phi(n)$, the number of integers between 1 and n coprime to n.

Any number n can be written as the product of its prime factors. Let

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}.$$

First, we look at a simple case, where $n = p_1^{\alpha_1} p_2^{\alpha_2}$. Our universal set S consists of the numbers between 1 and n, so $|S| = n$. We can construct two subsets, A_1 consisting of all the numbers between 1 and n that have p_1 as a factor and A_2 consisting of all the numbers between 1 and n that have p_2 as a factor. To find the number of integers between 1 and n coprime to n, we can subtract the number of integers that are not coprime to n (the ones that share a common divisor, be it p_1, p_2, or both, p_1p_2). By the PIE$_2$,

$$|S - A_1 \cap A_2| = |S| - |A_1| - |A_2| + |A_1 \cap A_2|.$$

The size of the set A_i is $\frac{n}{p_i}$.

Further, $A_1 \cap A_2$ will consist of all the numbers between 1 and n that are divisible by both p_1 and p_2, and $|A_1 \cap A_2| = \frac{n}{p_1p_2}$.

Thus,

$$\phi(n) = n - \frac{n}{p_1} - \frac{n}{p_2} + \frac{n}{p_1p_2}.$$

Factoring gives:

$$\begin{aligned}\phi(n) &= n\big(1 - \tfrac{1}{p_1} - \tfrac{1}{p_2} + \tfrac{1}{p_1p_2}\big) \\ &= n\big(1 - \tfrac{1}{p_1}\big)\big(1 - \tfrac{1}{p_2}\big).\end{aligned}$$

Similarly, when n has three prime factors, p_1, p_2, and p_3, let S be the set of numbers between 1 and n, let A_1 be the numbers with p_1 as a factor, let A_2 be the numbers with p_2 as a factor, and let A_3 be the numbers with p_2 as a factor. Then, the intersection of the two sets A_2 and A_3 is the set of numbers with p_2p_3 as a factor, and the intersection of all three sets is the set of numbers with $p_1p_2p_3$ as a factor. So by the PIE$_3$,

$$\begin{aligned}\phi(n) &= n\Big(1 - \tfrac{1}{p_1} - \tfrac{1}{p_2} - \tfrac{1}{p_3} + \tfrac{1}{p_1p_2} + \tfrac{1}{p_1p_3} + \tfrac{1}{p_2p_3} - \tfrac{1}{p_1p_2p_3}\Big) \\ &= n\Big(1 - \tfrac{1}{p_1}\Big)\Big(1 - \tfrac{1}{p_2}\Big)\Big(1 - \tfrac{1}{p_3}\Big).\end{aligned}$$

This can be generalized for $n = p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_r^{\alpha_r}$. Let $S = \{1, 2, ..., r\}$, and let $A_i = \{k \le r|\ p_i \text{ divides } k\}$. Thus,

$$\begin{aligned}\phi(n) &= |S| - |A_1 \cup A_2 \cup \cdots \cup A_n| \\ &= n\Big(1 - \tfrac{1}{p_1}\Big)\Big(1 - \tfrac{1}{p_2}\Big)\cdots\Big(1 - \tfrac{1}{p_r}\Big),\end{aligned}$$

or, using product notation:

$$\phi(n) = n\prod_{i=0}^{r}\Big(1 - \frac{1}{p_i}\Big).$$

4. Principle of Inclusion and Exclusion the Second

The former definition of the PIE, while straightforward, has the rather serious drawback that it rarely fits on one LaTeXpage. An alternate definition given in Lovász' book, is much shorter.

As before, given some finite set S, let $A_1, A_2, ..., A_n$ be subsets of S. For any I that is a subset of $\{1, 2, ..., n\}$, let

$$A_I = \bigcap_{i \in I} A_i$$

$$A_\emptyset = S.$$

So, the set A_I consists of the intersection of all the subsets A_i, where i is an element of I. Some examples of sets under the new definition are as follows:

$$A_{\{1\}} = A_1$$
$$A_{\{1,2\}} = A_1 \cap A_2$$
$$A_{\{5,12,22\}} = A_5 \cap A_{12} \cap A_{22}.$$

To define the PIE for these intersections, first assume S has three subsets: A_1, A_2, and A_3. Recall the PIE_3 states that:

$$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|.$$

Each of the sets on the right side of the equation can be described by A_I, where $I \subseteq \{1, 2, 3\}$.

$$\begin{array}{lll} A_1 = A_{\{1\}} & A_1 \cap A_2 = A_{\{1,2\}} & A_1 \cap A_2 \cap A_3 = A_{\{1,2,3\}} \\ A_2 = A_{\{2\}} & A_1 \cap A_3 = A_{\{1,2\}} & \\ A_3 = A_{\{3\}} & A_2 \cap A_3 = A_{\{2,3\}} & \end{array}$$

By the PIE_3, we know

$$\begin{aligned} |A_1 \cup A_2 \cup A_3| = \; & |A_{\{1\}}| + |A_{\{2\}}| + |A_{\{3\}}| \\ & -|A_{\{1,2\}}| - |A_{\{1,3\}}| - |A_{\{2,3\}}| \\ & +|A_{\{1,2,3\}}|. \end{aligned}$$

While this doesn't look any simpler than the previous definition, we can group each of the new subsets according to the size of I. The size of the union of A_1, A_2 and A_3 is:

$$|A_1 \cup A_2 \cup A_3| = \sum_{|I|=1} |A_I| - \sum_{|I|=2} |A_I| + \sum_{|I|=3} |A_I|. \tag{1}$$

Often, the PIE is applied to a set to calculate how much of the set S lies outside of its subsets. So, we'd like to be able to apply this definition to the problem $S - (A_1 \cup A_2 \cup A_3)$. Since we've defined

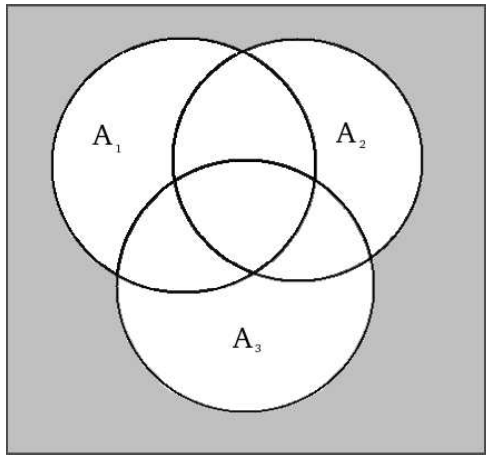

FIGURE 7. The shaded region is the set $S - (A_1 \cup A_2 \cup A_3)$. The size of this set can be written either as

$$|S - (A_1 \cup A_2 \cup A_3)| \text{ or as } |S| - |A_1 \cup A_2 \cup A_3|$$

.

$A_\emptyset = S$, this won't be exceptionally difficult. By inspecting Figure 4, we can see that .

$$|S - (A_1 \cup A_2 \cup A_3)| = |S| - |A_1 \cup A_2 \cup A_3|. \tag{2}$$

Writing Equation 2 in summation form is even simpler than in the above Equation 1, since the inclusion of the set $A_\emptyset$ guarantees that we're using all possible subsets of $\{1, 2, 3\}$. Thus,

$$|S - A_1 \cup A_2 \cup A_3| = \sum_{I \subseteq \{1,2,3\}} (-1)^{|I|} |A_I|. \tag{3}$$

Equation 3 can be generalized quite easily to a set S with n subsets $A_1, A_2, ...A_n$,

$$|S - A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{I \subseteq \{1,2,...,n\}} (-1)^{|I|} |A_I|. \tag{4}$$

4.1. **Example.** We can use the new PIE to calculate the number of onto functions with domain of size k and range or size n. Let the domain be the set $\{1, 2, ..., k\}$ and the codomain be the set $\{1, 2, ..., n\}$.

First, let the set S refer to the total number of functions with the domain and codomain assigned above. Since there are k elements in the domain and n members of the codomain, there are n^k possible functions between them.

In order to use the PIE, we need to assign the elements of subsets. Let A_i be the set of functions with domain $\{1, 2, ..., k\}$ and codomain the set $\{1, 2, ..., n\}$ without the element i, in other words: $\{1, 2, .., i-1, i+1, ..., n\}$. So, A_5 would be the functions with codomain $\{1, 2, 3, 4, 6, ...n\}$. Then the set $A_2 \cap A_4$ is the set of functions with codomain $\{1, 3, 5, 6, ..., n\}$.

To find the number of onto functions, we'll subtract the number of functions that miss at least one element of the set $\{1, 2, ..., n\}$, so we want to find a formula for

$$|S - (A_1 \cup A_2 \cup \cdots \cup A_n)|.$$

Recall Equation 4 gives the following formula for the PIE:

$$|S - (A_1 \cup A_2 \cup \cdots \cup A_n)| = \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} |A_I|.$$

Each of the sets A_1, A_2,..., A_n will have the same size, since there are equal numbers of sets with the same size domain and codomain, regardless of which element of $\{1, 2, ..., n\}$ is missing. Thus, every set A_I with $|I| = 1$ has the same size. There will be n sets with $|I| = 1$, and each set will have size $(n-1)^k$, since there are n possible choices to remove an element of $\{1, 2, ..., n\}$, and there are $(n-1)^k$ functions with domain of size k and codomain of size $n-1$.

Similarly, when we calculate the size of the sets A_I when $|I| = 2$, there will be $\binom{n}{2}$ sets, since there are $\binom{n}{2}$ ways to choose 2 elements from the set of size n. These functions all have domain of size k and codomain of size $n-2$. Each set will have size $(n-2)^k$.

Thus, for A_I when $|I| = i$, we would have $\binom{n}{i}$ sets of size $(n-i)^k$.

Thus, the number m of onto functions will be given by

$$m = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)^k.$$

Because $\binom{n}{i} = \binom{n}{n-i}$, we can revise the sum to show

$$m = \sum_{i=0}^{n} (-1)^i \binom{n}{n-i} (n-i)^k = \sum_{i=0}^{n} (-1)^{n+i} \binom{n}{i} i^k. \tag{5}$$

If $k < n$, then obviously $m = 0$, since no onto functions can exist. If $n = k$, then any onto funciton is one-to-one and $m = n!$. Thus,

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} i^k = \begin{cases} 0 & 0 \le k < n \\ (-1)^n n! & k = n \end{cases} \tag{6}$$

Equation 6 is an identity given by Lovász. For the remaining case, when $k > n$, we have that the codomain of the function is greater than its domain. Though it's not easy to use, Equation 6 is the simplest way to count these onto functions.

5. Möbius Functions

Before we go on, we define what it means for a set to be partially ordered by some relation. Given a set X, any subset of the product $\mathcal{R} \subseteq X \times X$ is called a relation. We sometimes write $a\mathcal{R}b$ to represent $(a, b) \in \mathcal{R}$. A relation is a partial ordering if it has all of the following three properties [1].

(1) Reflexivity: Given $x \in X$, we have $x\mathcal{R}x$ or $(x, x) \in \mathcal{R}$. That is, every element is related to itself.
(2) Antisymmetry: if $x\mathcal{R}y$ and $y\mathcal{R}x$, then $x = y$. If any two elements are each related to each other, then they must be equal.
(3) Transitivity: if $x\mathcal{R}y$ and $y\mathcal{R}z$, then $x\mathcal{R}z$.

Given any relation, we can determine whether it is a partial ordering by checking it against the above properties. Let's look at an example on the non-negative integers. Suppose $x\mathcal{R}y$ whenever $x|y$, so $\mathcal{R}$ contains all the pairs (x, y) where $x|y$, or whenever there exists some integer k such that $xk = y$.

(1) Reflexivity: Since any number divides itself, $x|x$, so $(x, x) \in \mathcal{R}$.
(2) Antisymmetry: Suppose $x|y$ and $y|x$. Thus there exist integers k and l such that $kx = y$ and $ly = x$. We can combine these equatinos to find that $klx = x$. Either $x = 0$, which is a trivial solution, or $kl = 1$. Since k and l are both non-negative integers, $k = l = 1$, so $x = y$.
(3) Transitivity: Suppose $x|y$ and $y|z$. Then there exist k and l such that $kx = y$ and $ly = z$. Thus, $klx = z$. Since k and l are both integers, kl will also be an integer, so $x|z$.

The **Möbius Function** of V, a finite set $\{x_1, ..., x_n\}$ partially ordered by the relation $\leq$, is the function μ defined on $V \times V$ such that

$$\begin{aligned} \mu(x,y) &= 0 \quad \text{if } x \not\leq y \\ \mu(x,x) &= 1 \\ \textstyle\sum_{x\leq y\leq z} \mu(x,y) &= 0 \quad (x < z) \end{aligned}$$

[2].

An easy way of examining relations is to think of them as lines connecting dots. An example is shown in Figure 8, which illustrates the Möbius Function of $V = \{a, b, c, d\}$. In this case, a dot x is related

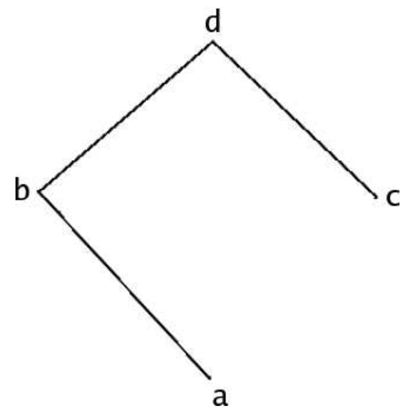

FIGURE 8. A partial ordering of the set $\{a, b, c, d\}$

to another dot y, or $x \leq y$, if they are connected and x is not above y. Therefore, every dot is related to itself and

$$a \leq b$$
$$a \leq d$$
$$b \leq d$$
$$c \leq d.$$

To determine the Möbius Function for Figure 8, we need to find $\mu(x, y)$ for every possible combination of x and y. From the definition of the Möbius Function, we know that for all $x \in V$ $\mu(x, x) = 1$, so

$$\mu(a, a) = \mu(b, b) = \mu(c, c) = \mu(d, d) = 1.$$

Also, $\mu(x, y)$ will be zero whenever x is not related to y. Since d is the maximum point, it isn't related to any other points, so

$$\mu(d, a) = \mu(d, b) = \mu(d, c) = 0.$$

Similarly, since b and c are both above a, we have $\mu(b, a) = \mu(c, a) = 0$. There is no line connecting b and c or a and c, so $\mu(a, c) = \mu(b, c) = \mu(c, b) = 0$.

We're now left with unknowns for $\mu(a, b)$, $\mu(a, d)$, $\mu(b, d)$, and $\mu(c, d)$. We first look at $\mu(a, b)$.

Recall from the definition of Möbius Functions that given any x and $z \in V$,

$$\sum_{x \leq y \leq z} \mu(x, y) = 0.$$

Since $a \leq a$ and $a \leq b$ and there is no other x such that $a \leq x \leq b$, we have $\mu(a, a) + \mu(a, b) = 0$. We know already that $\mu(a, a) = 1$, so

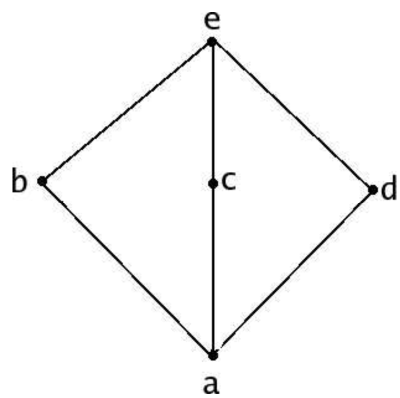

FIGURE 9. A partial ordering of the set $\{a, b, c, d, e\}$

$\mu(a, b) = -1$. We can calculate $\mu(b, d)$ and $\mu(c, d)$ the same way. We know $\mu(b, b) + \mu(b, d) = 0$ and $\mu(b, b) = 1$, so $\mu(b, d) = -1$. Also, $\mu(c, c) + \mu(c, d) = 0$ and $\mu(c, c) = 1$, so $\mu(c, d) = -1$. Calculation of $\mu(a, d)$ is slightly more complicated. There are three elements x that have $a \leq x \leq d$:

$$a \leq a \leq d$$

$$a \leq b \leq d$$

$$a \leq d \leq d$$

Since

$$0 = \sum_{a \leq x \leq d} \mu(a, x),$$

we have

$$\begin{array}{rcccccc} 0 & = & \mu(a,a) & + & \mu(a,b) & + & \mu(a,b) \\ & = & 1 & + & (-1) & + & \mu(a,d) \end{array}$$

Therefore, $\mu(a, d) = 0$.

This method of piecewise calculating μ for each combination of elements in a set can be used for more complicated relations as well. In Figure 9,

$$\mu(a, b) = \mu(a, c) = \mu(a, d) = \mu(b, e) = \mu(c, e) = \mu(d, e) = -1.$$

In this figure, $\mu(a, e)$ is the most interesting to calculate, since there will be five terms included in the sum.

Because a is related to every other point, and every point is also related to e,

$$\begin{gathered} a \leq a \leq e \\ a \leq b \leq e \\ a \leq c \leq e \\ a \leq d \leq e \\ a \leq e \leq e, \end{gathered}$$

Thus,

$$\sum_{a \leq x \leq e} \mu(a,x) = \mu(a,a) + \mu(a,b) + \mu(a,c) + \mu(a,d) + \mu(a,e) = 0$$

Recall that $\mu(a,a) = 1$. Note that b, c, and d are each only one step above a; that is, there are no other points x for which $a \leq x \leq b$. Therefore, $\mu(a,b) = \mu(a,c) = \mu(a,d) = -1$. So,

$$\mu(a,e) = 0 - 1 - 3(-1) = 2.$$

Even for the most complicated relations, determining $\mu(x,y)$ for any x and y is a simple process. The relation between x and y can be broken down into small steps, each of which we can calculate.

Let's examine the Möbius Functions of some specific relations on various sets V.

5.1. **Subsets.** First let's consider the subsets of some set S, which are ordered by inclusion. Since they're ordered by inclusion, we say $A \leq B$ if and only if $A \subseteq B$. We'll begin with a short example. Given the set $S = \{1,2,3\}$, we can construct a diagram as shown in Figure 10. Whenever two sets are connected, that means the lower set is a subset of the higher one. Note that there is a top set, the maximum, located at the top, back, right corner of the cube; each subset is related to the set S. Also, the empty set, located at the front, bottom, left corner, is a subset of every other set.

The Möbius Function in this case will be

$$\begin{aligned} \mu(A,A) = {} & 1 \\ \mu(A,B) = {} & 0 \quad \text{if } A \not\subseteq B \\ \textstyle\sum_{A \subseteq C \subseteq B} \mu(A,C) = {} & 0 \quad \text{if } A \subseteq B \end{aligned}$$

The Möbius Functions of each subset with itself will be equal to 1, and then when $|A| - |B| = 1$, we have $\mu(A,B) = -1$. Now we examine a pair of subsets whose sizes differ by 2. We'll find $\mu(A,B)$ where $A = \{1\}$ and $B = \{1,2,3\}$. According to the third part of the formula for constructing a Möbius Function,

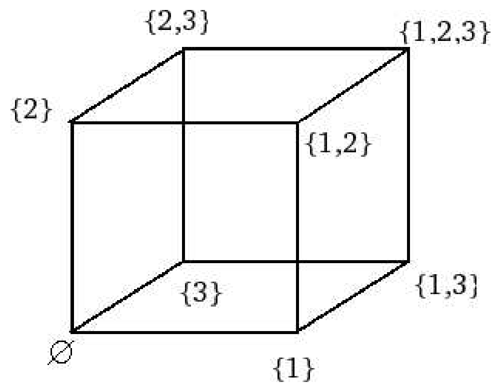

FIGURE 10. The partial ordering with respect to set inclusion of the set $S = \{1, 2, 3\}$

$$\sum_{\{1\} \subseteq X \subseteq \{1,2,3\}} \mu(\{1\}, X) = 0.$$

Thus,

$$\mu(\{1\}, \{1\}) + \mu(\{1\}, \{1, 2\}) + \mu(\{1\}, \{1, 3\}) + \mu(\{1\}, \{1, 2, 3\}) = 0.$$

We know $\mu(\{1\}, \{1\}) = 1$, and $\mu(\{1\}, \{1, 2\}) = \mu(\{1\}, \{1, 3\}) = -1$. Thus, $\mu(\{1\}, \{1, 2, 3\}) = 1$. Therefore, the Möbius Function of any two subsets with a difference in size of 2 is 1. It now appears that we have a formula for the Möbius Function of subsets ordered by inclusion:

$$\begin{array}{rcl} \mu(A, A) = & 1 & \\ \mu(A, B) = & 0 & \text{if } A \not\subseteq B \\ \mu(A, B) = & (-1)^{|B|-|A|} & \text{if } A \subseteq B \end{array}$$

We can prove this result by induction. Let n be the difference in size between sets A and B. So, $n = |B| - |A|$. We want to show that $\mu(A, B) = (-1)^n$. For the base case, let $n = 1$. Since set B has only one element more than B, there are no other sets C such that $A \subseteq C \subseteq B$. Therefore,

$$\mu(A, A) + \mu(A, B) = 0.$$

Since $\mu(A, A) = 1$, we find that $\mu(A, B) = -1$, which is the same as $(-1)^1 = (-1)^n$.

Assume that our formula holds true for all $k < n$. So, assume that when $|B| - |A| = k$, we have $\mu(A, B) = (-1)^k$. We want to show that when $|B| - |A| = n$, we have $\mu(A, B) = (-1)^n$. Since $A \subset B$, we know B consists of all the elements of A as well as n extra elements.

Recall

$$\sum_{A\subseteq C\subseteq B} \mu(A,C) = 0 \tag{7}$$

There exist sets C such that for all $x < n$, $|C| = |A| + x$. So, for $x = 1$, C has one element more than A, and that one element is one of the extra n elements in B. There are therefore $\binom{n}{1}$, or n, sets C. Similarly, for $|C| = |A| + 2$, C consists of A and two other elements that are two of the n extra elements in B, so there are $\binom{n}{2}$ sets C.

So, for C_i defined as any set with $|C_i| = |A| + i$, we could rewrite Equation 7:

$$\begin{aligned} 0 &= \textstyle\sum_{A\subseteq C\subseteq B} \mu(A,C) \\ &= \mu(A,A) + \binom{n}{1}\mu(A,C_1) + \binom{n}{2}\mu(A,C_2) + \cdots + \binom{n}{n}\mu(A,B) \end{aligned} \tag{8}$$

Because of our initial assumption, the $\mu(A,C)$ is known for all C smaller than B, and $\mu(A,C_i) = (-1)^i$.

Thus,

(9)

$$0 = \sum_{A\subseteq C\subseteq B} \mu(A,C) = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} + \cdots + (-1)^{n-1}\binom{n}{n-1} + \binom{n}{n}\cdot\mu(A,B)$$

The numbers $\binom{n}{0}$ $\binom{n}{1}$ $\cdots$ $\binom{n}{n}$ form the n^{th} row of Pascal's Triangle, which also gives the coefficients for binomial expansion, as follows:

$$(a+b)^n = a^n + na^{n-1}b + \binom{n}{2}a^{n-1}b^2 + \cdots \binom{n}{n}b^n \tag{10}$$

If we choose $a = 1$ and $b = 1$, then Equations 10 and 11 are the same, and we have

$$0 = (1-1)^n = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots + (-1)^n\binom{n}{n}.$$

Since Equations 9 and 10 both add to 0, and every term up until the last is also equal, it follows that

$$\binom{n}{n}\cdot\mu(A,B) = \binom{n}{n}(-1)^n.$$

Therefore, $\mu(A,B) = (-1)^n$, and so for all $A \subseteq B$, where $n = |B| - |A|$, we have $\mu(A,B) = (-1)^n$.

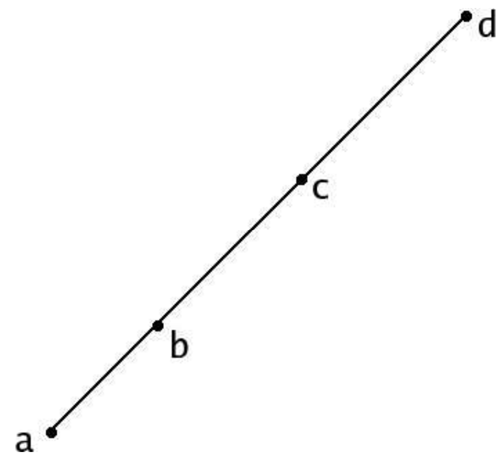

FIGURE 11. An Arborescence

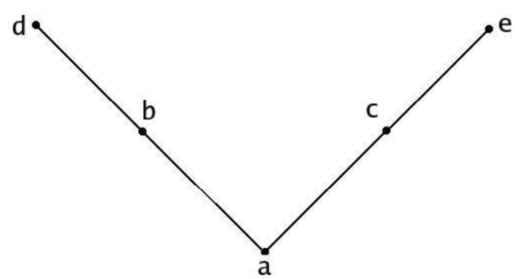

FIGURE 12. Another Arborescence

5.2. **Arborescences.** An arborescence is a partial ordering with a least element where there is a unique path between any two elements.

The set in Figure 8 is not an arborescence because it has two minimal elements, a and c. Figure 9 also doesn't demonstrate an arborescence. Even though a is a least element, there are three paths from a to e. Figures 11 and 12 both represent arborescences. In each case, a is the least element and there is only one possible path to any of the other points.

To define the Möbius Function for arborescences, let's start with a, the least element shown in Figure 11. We know $\mu(a, a) = 1$. The next point, b, will have $\mu(a, b) = -1$, so that $\mu(a, a) + \mu(a, b) = 0$. Since there's only one possible path between any two points,

$$\mu(a, c) = \mu(a, d) = 0,$$

since it must be true both that

$$\mu(a, a) + \mu(a, b) + \mu(a, c) = 0$$

and

$$\mu(a,a)+\mu(a,b)+\mu(a,c)+\mu(a,d)=0.$$

If we start at b, we find that $\mu(b,b)=1$, then $\mu(b,c)=-1$, and it follows that $\mu(c,d)=0$.

Similarly, in Figure 12, $\mu(a,b)=\mu(a,c)=-1$, since both b and c are only one step away from a. Then $\mu(a,d)=\mu(a,e)=0$, since in both cases there is only one possible way to get from a to d or from a to e, and so

$$\mu(a,a)+\mu(a,b)+\mu(a,d)=0$$

and

$$\mu(a,a)+\mu(a,c)+\mu(a,e)=0.$$

Therefore, for an arborescence V that contains some elements x, y, and z, if we're finding $\mu(x,z)$ and there is no y other than x and z such that $x\leq y\leq z$, then $\mu(x,z)=-1$. Since every path is unique, every z after that, where there is at least one y such that $x\leq y\leq z$ will have $\mu(x,z)=0$.

5.3. **Integers.** Now let's examine the Möbius Function for the set of integers $1,...,n$, and we say $x\leq y$ whenever $x|y$.

Since there are many combinations of integers a and b such that $a|b$, it would be nice if we could find some way to simplify the problem. We're going to show that $\mu(a,b)=\mu(1,\frac{b}{a})$ by induction. Given any a, we know $\mu(1,1)=\mu(a,a)$. Now suppose that $\mu(1,\frac{c}{a})=\mu(a,c)$ for all $c<b$ such that $a|c$ and $c|b$.

Note that

$$\mu(a,b)=-\sum_{a\leq c<b}\mu(a,c).$$

Thus, by the induction hypothesis,

$$\mu(a,b)=-\sum_{a\leq c<b}\mu(1,\frac{c}{a}).$$

Instead of summing over all c such that $a|c$ and $c|b$, we can sum over all $\frac{c}{a}$ such that $1|\frac{c}{a}$ and $\frac{c}{a}|\frac{b}{a}$. So we now have

$$\begin{aligned}\mu(a,b) &= -\textstyle\sum_{1\leq\frac{c}{a}<\frac{b}{a}}\mu(1,\frac{c}{a})\\ &= \mu(1,\tfrac{b}{a}).\end{aligned}$$

Let $\frac{b}{a}=n$. Since n is some integer, it can be written as the product of its prime factors; that is, $n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_r^{\alpha_r}$.

To find the formula for $\mu(1,n)$, let's examine separately the following cases

(1) n is a product of distinct primes

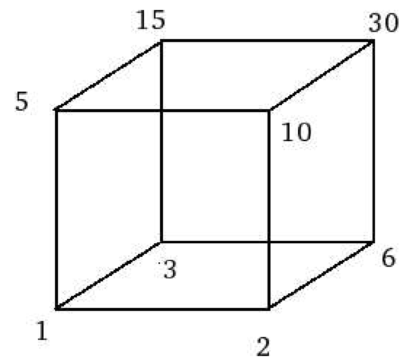

Figure 13. The partial ordering with respect to division of $n = 30$

(2) n has at least one repeated factor

5.3.1. *Case(1)* $n = p_1 \cdot p_2 \cdots p_r$. Figure 13 shows a simple example of the partial ordering of $n = 2 \cdot 3 \cdot 5$.

Notice that the structure of the partial ordering in Figure 13 is identical to that of the subsets examined in the previous section, shown in Figure 10. For integers, 1 takes the place of the emptyset, and n takes the place of A, and all the multiples of the prime factors of n correspond to the remaining subsets of A. This result extends to any r since every integer x such that $x|p_1p_2...p_r$ corresponds to a subset of $\{1, 2, ..., r\}$. Thus, the same result follows for $\mu(1, n)$ as for $\mu(A, B)$. When n has r distinct prime factors, $\mu(1, n) = (-1)^r$.

5.3.2. *Case(2):* $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$*, where some* $\alpha_i \neq 1$. We would like to use induction to show that $\mu(1, n) = 0$.

Assume for all numbers m smaller than n, where $m = p_1^{\alpha_1} \cdots p_q^{\alpha_q}$

$$\mu(1, m) = \begin{cases} (-1)^q & \text{if every } \alpha = 1 \\ 0 & \text{if there exists some } \alpha \text{ such that } \alpha \neq 1 \end{cases}$$

Let $n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_r^{\alpha_r}$. We can write $\mu(1, n)$ as a sum of the Möbius Functions of all of its factors:

$$\mu(1, n) = -\sum_{1 \leq a < n} \mu(1, a)$$

Of the a's, some will be products of distinct prime factors of n, and the rest will have repeated prime factors. All of the a's with repeated prime factors will have $\mu(1, a) = 0$. Those with distinct prime factors are all divisors of the one a_r which is the product of all the prime factors of n, where $a_r = p_1 \cdot p_2 \cdots p_r$. Since all the other a's with distinct prime factors divide a_r, we know that

$$\sum_{1 \le a \le a_r} \mu(1,a) = 0.$$

Thus, the Möbius functions of all the divisors of n with distinct prime factors will add to zero, and the Möbius functions of all the divisors of n with repeated prime factors are 0. Therefore, since

$$\mu(1,n) = -\sum_{1 \le a \le n} \mu(1,a),$$

$$\mu(1,n) = 0.$$

So, now we have a formula for the Möbius functions of any two integers:

$$\begin{array}{lcll} \mu(x,y) & = & 0 & \text{if } x \text{ does not divide } y \\ \mu(x,x) & = & 1 & \\ \mu(x,y) & = & \mu(1,\frac{y}{x}) & \\ \mu(1,\frac{y}{x}) & = & (-1)^r & \text{if } \frac{y}{x} = p_1 \cdots p_r \\ \mu(1,\frac{y}{x}) & = & 0 & \text{if } \frac{y}{x} = p_1^{\alpha_1} \cdots p_r^{\alpha_r} \text{ with some } \alpha > 1. \end{array}$$

6. Upside-Down Möbius Function

Before we can examine the Möbius Inversion Formula, there's one very important Möbius Function to consider. The Upside-Down Möbius Function acts on V ordered by $\le^*$, where $x \le^* y$ whenever $y \le x$. We define the Upside-Down Möbius Function as follows:

$$\begin{array}{rcll} \mu^*(x,y) & = & 0 & \text{if } x \not\le^* y \\ \mu^*(x,x) & = & 1 & \\ \sum_{x \le^* y \le^* z} \mu^*(x,y) & = & 0 & (x <^* z) \end{array}$$

This form only illustrates that the Upside-Down Möbius Function is indeed a Möbius Function. To relate it to the previous sections, we'd like to show that $\mu^*(x,y) = \mu(y,x)$.

6.0.3. *Example.* To begin, let's look at an example. We'll calculate the Möbius Function and the Upside-Down Möbius Function acting on the set shown in Figure 14. As was shown earlier, $\mu(a,d)$ can be calculated by adding up the Möbius Functions of all the elements below d.

Recall that

$$\sum_{a \le x \le d} \mu(a,x) = 0.$$

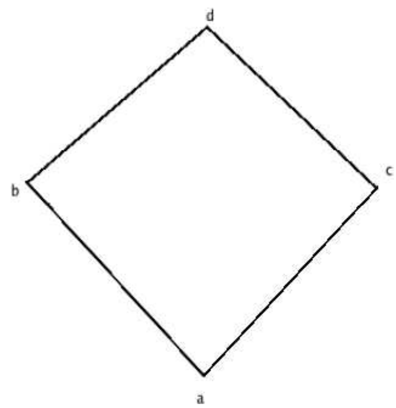

FIGURE 14. A relation on the set $\{a, b, c, d\}$

So, $\mu(a, a) + \mu(a, b) + \mu(a, c) + \mu(a, d) = 0$. To obtain a formula strictly for $\mu(a, d)$, we subtract everything except $\mu(a, d)$ to get

$$\begin{aligned} \mu(a,d) = & -(\mu(a,a) + \mu(a,b) + \mu(a,c)) \\ = & -\textstyle\sum_{a \le x < d} \mu(a,x). \end{aligned}$$

Because b and c are each only one step away from a, we know $\mu(a, b) = \mu(a, c) = -1$. Thus, simple addition gives $\mu(a, d) = 1$.

We can use exactly the same process to calculate the Upside-Down Möbius Function for the same set.

Our definition for the Upside-Down Möbius Function stated

$$\sum_{d \le^* x \le^* a} \mu^*(d, x) = 0.$$

Since $d \le^* x$ is equivalent to saying $x \le d$, we can rewrite the bounds in the sum:

$$\sum_{a \le x \le d} \mu^*(d, x) = 0.$$

This is equivalent to saying

$$\mu^*(d, a) + \mu^*(d, b) + \mu^*(d, c) + \mu^*(d, d) = 0.$$

Since we're mostly interested in the Upside-Down Möbius Function between d and a, we can rewrite this as a formula for $\mu^*(d, a)$:

$$\mu^*(d, a) = -(\mu^*(d, b) + \mu^*(d, c) + \mu^*(d, d)),$$

or equivalently

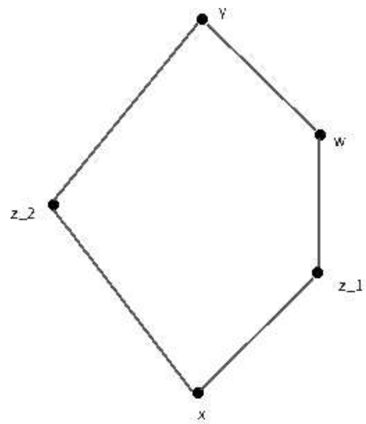

FIGURE 15. A slightly more compolicated relation on the set $\{x, z_1, z_2, w, y\}$.

$$\mu^*(d, a) = \sum_{a<x\leq d} \mu^*(d, x).$$

The Upside-Down Möbius Function is defined in the same was as the regular Möbius Function, so we can calculate $\mu^*(d, c)$ and $\mu^*(d, b)$ in the same way as we would if they were normal. The only difference is here we're adding down instead of up the relation. Since b and c are only one step away from d, then $\mu^*(d, b) = \mu^*(d, c) = -1$, so $\mu^*(d, a) = 1$. In this case, at least, it is true that $\mu(a, d) = \mu^*(d, a)$.

Recall that

$$\mu(a, d) = -(\mu(a, a) + \mu(a, b) + \mu(a, c)).$$

Both $\mu(a, b)$ and $\mu(a, c)$ can also be expressed as sums of Möbius Functions of elements that are 'below' them. In this case, only a is below either of these, so $\mu(a, b) = \mu(a, c) = -\mu(a, a)$. Thus,

$$\begin{aligned} \mu(a, d) = & -(\mu(a, a) - (\mu(a, a)) - (\mu(a, a)) \\ = & -(1 - (1) - (1)) \\ = & (-1)^1 + (-1)^2 + (-1)^2 \\ = & 1. \end{aligned}$$

6.0.4. *Another Example.* Let's look at a more complicated relation, as shown in Figure 15.

In Figure 15, the points are now labeled according to what "level" they are. The lowest level consists of x, and the top level is only y. Both z_1 and z_2 are exactly one step up from x, and then w is a step above z_1. We're going to examine $\mu(x, y)$.

We found before that the Möbius Function of any two objects can be expressed as a negative sum of the Möbius Functions of the objects between them. Thus, we can write $\mu(x, y)$ as

$$\mu(x, y) = -(\mu(x, x) + \mu(x, z_1) + \mu(x, z_2) + \mu(x, w)).$$

Because z_1 and z_2 are only one step away from x, each can be expressed as $-(\mu(x, x))$. Because there is an intermediate step in between x and w, $\mu(x, w)$ will be the sum $-(\mu(x, x) + \mu(x, z_2)$. Then, $\mu(x, z_2)$ is further reduced to $-\mu(x, x)$. In this way, $\mu(x, y)$ is simplified to a sum of powers of (-1) as shown below.

Thus,

$$\begin{aligned}
\mu(x, y) &= -(\mu(x, x) - (\mu(x, x)) - (\mu(x, x)) - (\mu(x, x) + \mu(x, z_2))) \\
&= -(\mu(x, x) - (\mu(x, x)) - (\mu(x, x)) - (\mu(x, x) + (-\mu(x, x)))) \\
&= -(1 - (1) - (1) - (1 - (1))) \\
&= (-1)^1 + (-1)^2 + (-1)^2 + (-1)^2 + (-1)^3 \\
&= (-1)^1 + 3(-1)^2 + (-1)^3 \\
&= 1.
\end{aligned}$$

Notice these numbers describe the possible paths between x and y. Since $x \leq y$, there is a path of length 1 that connects x and y. There are three possible ways to go from x to y with one intermediate step:

$$\begin{array}{ccccc}
x & \leq & z_1 & \leq & y \\
x & \leq & z_2 & \leq & y \\
x & \leq & w & \leq & y
\end{array}$$

Thus, there are three occurences of $(-1)^2$. There is only one possible way to take three steps going from x to y:

$$x \leq z_2 \leq w \leq y,$$

and we have only one $(-1)^3$.

So, we can also express $\mu(x, y)$ as follows:

$$\mu(x, y) = \sum_{x \leq q_1, q_2, \ldots, q_i \leq y} (-1)^{i+1},$$

where q_i describes the intermediate steps in between x and y; $(-1)^{i+1}$ is summed over all possible paths between x and y, and i is the length of the path.

Since there will be the same number and size of paths going from y to x as there were from x to y, we can say that

$$\mu^*(y, x) = \sum_{x \leq q_1, q_2, \ldots, q_i \leq y} (-1)^{i+1}.$$

So, $\mu(x,y) = \mu^*(y,x)$.

These same concepts can be used for any relation and arbitrary x and y so that, in general,

$$\mu(x,y) = \mu^*(y,x).$$

7. Möbius Inversion Formula

Now, we finally have enough background to introduce the Möbius Inversion Formula: Given any function $f(x)$, if $g(x)$ is defined such that

$$g(x) = \sum_{z \leq x} f(z),$$

then

$$f(x) = \sum_{z \leq x} g(z)\mu(z,x). \tag{11}$$

This result can be proven using induction.

First, note that

$$f(x) = \sum_{z \leq x} f(z) - \sum_{z < x} f(z). \tag{12}$$

Let $g(x) = \sum_{z \leq x} f(z)$. The first term in Equation 12 is therefore simply $g(x)$.

For the first case, there will be no z such that $z < x$, so $f(x) = g(x)$, and since $\mu(x,x) = 1$, $f(x) = g(x)\mu(x,x)$.

We can assume that for any element w less than the current x, $f(w) = \sum_{z \leq w} g(z)\mu(z,w)$. Thus,

$$f(x) = g(x) - \sum_{z<x} \sum_{w \leq z} g(w)\mu(w,z)$$

Any given $g(w)$ will appear with each z where $w \leq z < x$, which gives

$$f(x) = g(x) - g(w)(-\mu(w,z_1) - \mu(w,z_2) - ... - \mu(w,w).$$

which is the same as

$$f(x) = g(x) - g(w)(-\sum_{z<x} \mu(w,z)$$

Recall that

$$\mu(x,y) = -\sum_{z<y} \mu(x,z).$$

The term in the sum associated with each w will be $g(w)\mu(w, x)$. The set of w's is the same as the set of z's, since the w's are defined as the elements such that $w \leq z < x$. Any z has at least one w, itself, and every w is a z, since $w < x$.

So, Equation 12 can be rewritten as:

$$f(x) = g(x) + \sum_{z<x} g(z)\mu(z, x). \tag{13}$$

Since $\mu(x, x) = 1$, we can multiply $g(x)$ by $\mu(x, x)$, which allows us to simplify Equation 13 to

$$f(x) = \sum_{z \leq x} g(z)\mu(z, x). \tag{14}$$

Note that Equations 14 and 11 are the same.

Therefore, when

$$g(x) = \sum_{z \leq x} f(z),$$

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

and we've proved the Möbius Inversion Formula.

8. Möbius PIE

Recall the Principle of Inclusion and Exclusion for a set S with subsets A_1, A_2, ..., A_n:

(15)

$$\begin{aligned} |S - A_1 \cup \cdots \cup A_n| = \quad |S| \quad & -(|A_1| + |A_2| + \cdots + |A_n|) \\ & +(|A_1 \cap A_2| + |A_1 \cap A_3| + \cdots + |A_{n-1} + A_n| \\ & -(|A_1 \cap A_2 \cap A_3| + \cdots + |A_{n-2} \cap A_{n-1} \cap A_n|, \end{aligned}$$

or, alternatively, if we define I as some subset of $\{1, 2, ..., n\}$ and let

$$A_I = \bigcap_{i \in I} A_i \qquad A_\emptyset = S,$$

then

$$|S - (A_1 \cup \cdots \cup A_n)| = \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|}|A_I|. \tag{16}$$

Earlier we demonstrated the Principle of Inclusion-Exclusion using Venn diagrams for small numbers of subsets, and we proved it using induction. Now that we're familiar with the Möbius Inversion Formula, we can use it to prove the PIE, thereby coming full circle.

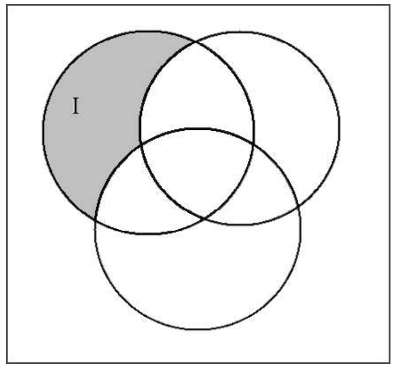

FIGURE 16. The shaded region has size $f(I)$

We'll be using the version of the PIE shown in Equation 16. To begin, consider subsets of $\{1, 2, ..., n\}$. Given two subsets, called I and J, we'll say $I \leq J$ whenever $J \subseteq I$. Recall the Möbius Function for subsets ordered by inclusion, which stated that if we have two sets A and B, and $A \subseteq B$, then $A \leq B$ and $\mu(A, B) = (-1)^{|B|-|A|}$. Notice that in our current example, we have that if $J \subseteq I$, then $I \leq J$. This is exactly backwards of the formula for the Möbius Function of subsets, which defines $\mu(J, I)$, since $J \subseteq I$. However, the Upside-Down Möbius Function guarantees that $\mu^*(I, J) = \mu(J, I)$. Thus, we can say $\mu(I, J) = (-1)^{|I|-|J|}$.

Now recall the Möbius Inversion formula, which stated that given any $f(x)$, if we define $g(x)$ such that

$$g(x) = \sum_{z \leq x} f(z)$$

then we will have

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

Let's define our $f(x)$ so that

$$f(I) = \left| \bigcap_{i \in I} A_i - \bigcup_{I \subset J} (\bigcap_{j \in J} A_j) \right|$$

or, equivalently,

$$f(I) = \left| A_I - \bigcup_{I \subset J} A_J \right|.$$

So, for some $I \subseteq \{1, 2, ..., n\}$, $f(I)$ is the size of the set consisting of elements that are in A_I but are not in any other subsets of S. Figure 8 shows the region whose size is $f(I)$.

Let's first look at the set S, or $A_\emptyset$.

$$f(\emptyset) = \left| A_\emptyset - \bigcup_{I \subseteq \{1,\dots,n\}} A_I \right| = \left| A_\emptyset - \bigcup_{i \in \{1,2,\dots,n\}} A_i \right|$$

So, if we look at the big set S, or $A_\emptyset$,

$$f(\emptyset) = \left| A_\emptyset - \bigcup_{I \subseteq \{1,\dots,n\}} A_I \right| = \left| A_\emptyset - \bigcup_{i \in \{1,2,\dots,n\}} A_i \right|.$$

This is the expression for $\big|S - (A_1 \cup \cdots \cup A_n)\big|$, so in order to prove the Principle of Inclusion-Exclusion, we need

$$f(\emptyset) = \sum_{I \subseteq \{1,2,\dots,n\}} |A_I|(-1)^{|I|}.$$

To construct $g(I)$ from the definition of the Möbius Inversion Formula, let

$$g(I) = \sum_{J \leq I} f(J).$$

Substituting the formula for $f(J)$ gives

$$g(I) = \sum_{J \leq I} \left| A_J - \bigcup_{K \subset J} A_K \right|.$$

If we separate out the first term I,

$$g(I) = \left| A_I - \bigcup_{I \subset J} A_J \right| + \sum_{I \subset J} \left| \big(A_J - \bigcup_{K \subset J} A_K\big) \right| \tag{17}$$

Not that the summation term in Equation 17 sums the sizes of the sets corresponding to each I that consist of elements that are only in I, and not in any subsets of I. Since we sum over all $I \subset J$, we can rewrite that sum as

$$\sum_{I \subset J} \left| \big(A_J - \bigcup_{K \subset J} A_K\big) \right| = \left| \bigcup_{I \subset J} A_J \right.$$

Thus, $g(I) = |A_I|$.

Remember from the original Möbius Inversion Formula,

$$f(I) = \sum_{J \leq I} g(J)\mu(J, I).$$

We know $\mu(J,I) = (-1)^{|J|-|I|}$, so

$$f(I) = \sum_{I \subseteq J} |A_J| \cdot (-1)^{|J|-|I|}.$$

Let's examine $f(\emptyset)$ again. This should give the size of the set containing all the elements of S that aren't any of the subsets $A_1, ..., A_n$, or $|S - A_1 \cup A_2 \cup \cdots \cup A_n|$. We find that

$$f(\emptyset) = \sum_{\emptyset \subseteq J} |A_J| \cdot (-1)^{|J|}$$

or

$$f(\emptyset) = \sum_{I \subseteq \{1,2,...,n\}} |A_I| \cdot (-1)^{|I|},$$

which looks pretty familiar, since it's the Principle of Inclusion-Exclusion!

References

[1] McEliece, Robert J., Robert B. Ash, Carol Ash. *Introduction to Discrete Mathematics* 1^{st} ed. New York: Random House, 1989.

[2] Lovász, L., *Combinatorial Problems and Exercises.* New York: North Holland Publishing Company, 1979.

[3] Schueller, Laura. Personal interview. 6 Apr.2008.