Pigeonhole Principle Mini Lecture

Here we will explore the pigeonhole principle which is a basic and relatively simple fact that can be used in clever ways to prove surprising results.

The pigeonhole principle is based on the following observation: If you have 10 letters each of which is going to be placed in one of 8 mailboxes, then at least one mailbox will have to more than 1 letter. In other words, some mailbox will have two or more. Another example is full moons, there are 13 full moons a year and there are 12 months, so some month must have at least two full moons. So the pigeonhole principle observes that for finite sets of objects anytime the number of inputs is larger than the number of outputs, some output must be 'hit' twice.

Theorem (The Basic Pigeonhole Principle) Suppose n objects are placed into k classes. If n > k, then *some* class contains more than one object.

Proof: We prove the contrapositive. Assume *no* class contains more than one object. Let a_1, a_2, \ldots, a_k be the number of objects in each class. Then $a_1 \leq 1, a_2 \leq 1, \ldots, a_k \leq 1$. Thus, the total number of objects, $n = \sum_{i=1}^{k} a_i = a_1 + \cdots + a_k \leq 1 + \cdots + 1 = k$. Hence, $n \leq k$. \Box .

Another way to say this is with functions. Suppose $f : A \to B$ is a function where A and B are finite. An equivalent way to state the basic pigeonhole principle is the following: If |A| > |B|, then f is not injective. (You essentially observed this in homework 4.45 because if A and B are finite and |A| > |B|, then f cannot be surjective which forces it to be not injective since we are working with finite sets).

To use the pigeonhole principle you need to be very clear and precise about how you are defining the *objects* and the *classes*. Once this is done, you just have to verify that the number of objects exceeds the number of classes, then you can conclude that some class contains more than one object. Here is a nontrivial example:

Theorem

Given any list of n positive integers, there exists two elements whose difference is divisible by n-1.

EXAMPLE: Before I do the prove, let's carefully work through an example: Let $S = \{1, 7, 9, 42, 106\}$ has n = 5 elements. This theorem claims there is some difference in here that is divisible by n - 1 = 4. Here are a few differences: 7 - 1 = 6, 9 - 7 = 2, 42 - 1 = 41, 42 - 7 = 35, so far we have not found anything divisible by 4, but this theorem says if we keep looking we will. And here is is 9 - 1 = 8 is divisible by 4. (Also 106 - 42 = 64 is divisible by 4).

The reason this works, for this particular example, is because when dividing by 4 there are 4 possible remainders (namely, 0, 1, 2, 3).

1 divided by 4 has remainder 1 (in other words $1 = 0 \cdot 4 + 1$)

7 divided by 4 has remainder 3 (in other words $7 = 1 \cdot 4 + 3$)

9 divided by 4 has remainder 1 (in other words $9 = 2 \cdot 4 + 1$)

42 divided by 4 has remainder 2 (in other words $42 = 10 \cdot 4 + 2$)

106 divided by 4 has reminder 2 (in other words $106 = 26 \cdot 4 + 2$)

If we do 9-1, then the remainders cancel and we get something divisible by 4, same for 106 - 42. Since there are 4 possible remainders and 5 numbers, the pigeonhole principle tells us that some remainder must occur more than once.

Now for the general proof.

Proof: Let a_1, a_2, \ldots, a_n be the list of n numbers. For each of these numbers divide by n-1 and get the remainder. That is, for $i = 1, 2, \ldots, n$, write $a_i = q_i(n-1) + r_i$ where the remainder $0 \le r_i \le n-2$ (the remainder theorem which we prove in chapter 6 guarantees this can always be done). Let the *objects* be these remainders, r_1, r_2, \ldots, r_n . So there are n objects. Let the classes be $0, 1, 2, \cdots n-2$ of which there are n-1. Each of the n remainders is in at one of the n-1 classes, so by the pigeonhole principle, some remainder occurs more than once.

Since we know at least two remainders are the same, let r_i and r_j be remainders that are the same. Then by definition, $a_i = q_i(n-1) + r_i$ and $a_j = q_j(n-1) + r_j$.

Subtracting gives $a_j - a_i = (q_j(n-1) + r_j) - (q_i(n-1) + r_i) = (q_j - q_i)(n-1).$

Hence n-1 divides the difference $a_j - a_i$. \Box

Here is another example you can think about at your next party.

Theorem

Suppose n people are at a party $(n \ge 2)$. There are at least two people at the party that have the same number of friends.

Proof: We let the *n* people be the objects and we sort them into *k* classes where *k* is the number of friends they have. The possible number of friends they have is k = 0, 1, 2, ..., n - 1.

Since we have n objects and n classes, the pigeonhole principle does not apply without proving more. But we can prove me, we will prove that either k = 0 or k = n - 1 must be empty.

If some person has k = 0 friends at the party, then it is impossible for some other person to have k = n - 1 friends (because then they would be friends with the zero friend person which we just said is not the case). By the same reasoning, if some person has k = n - 1 friends (meaning they are friends with everyone at the party), then no one can have k = 0 friends at the party.

Hence, one of the classes k = 0 or k = n - 1 must be empty.

Thus, there are only n-1 nonempty classes and there are n objects. By the pigeonhole principle, some two people have the same number of friends. \Box

We can extract a big more information by looking at exactly how much bigger n is than k. For example, if there are n = 7 objects and k = 3 classes, then some class much be hit more than **two** times. If you don't believe this then try it on the following set of 7 objects $\{A, B, C, D, E, F, G\}$, no matter how you break this into three groups, at least one group will always have at least two objects. This is because n = 7 if more than twice k = 3.

Using this observation, the pigeonhole principle is usually stated in a slightly more general way.

Theorem (The General Pigeonhole Principle) Suppose n objects are placed into k classes. If n > mk, then *some* class contains more than m objects.

Proof: We prove the contrapositive. Assume no class contains more than m objects. Let a_1, a_2, \ldots, a_k be the number of objects in each class. Then $a_1 \leq m, a_2 \leq m, \ldots, a_k \leq m$. Thus, the total number of objects, $n = \sum_{i=1}^k a_i = a_1 + \cdots + a_k \leq m + \cdots + m = km$. Hence, $n \leq km$. \Box .

For example, if n = 16 letters are to be delivered to k = 3 mailboxes, then since 16 is bigger than 5 times 3, some mailbox will get more than 5 letters.