

# A Cute Lemma from Ramsey Theory

Ramsey theory is a branch of combinatorics with many surprising results. For the most part, we normally think of Ramsey theory as a field that deals with extremely large sets, if not infinite sets. As a computer scientist (in training), I tend to be interested in a very specific class of functions though: specifically functions of the form  $f : X \rightarrow Y$  where  $X, Y$  are *discrete* and *finite* sets. For example,  $\{0, 1\}^n \rightarrow \{0, 1\}$  is a function that takes in a binary string of length  $n$  and outputs a single bit.

Though in this quick post, we will be discussing, **exponentially sized** sets, we will get a really interesting and *unexpected* result about a very common class of functions!

## The Class of Functions $F(x, y)$

Let  $X, Y$  be finite sets with  $|X| \geq 2^{k+2}$  and  $|Y| \geq 3$ . We will let  $F : X \times X \rightarrow Y$  be an arbitrary function as long as  $F(x, y) = F(y, x)$  for all  $x, y \in X$ . We call these class of functions *symmetric* functions. (TODO: check this?)

**Definition** (Symmetric Function): A function  $F : X \times X \rightarrow Y$  is called *symmetric* if  $F(x, y) = F(y, x)$  for all  $x, y \in X$ .

## Cute Little Lemma for $F(x, y)$

**Theorem:** For all  $F : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{Y}$  where  $F$  is a symmetric function and  $|\mathcal{X}| \geq 2^{k+2}, |\mathcal{Y}| \geq 3$ , there exists a set  $\mathcal{S} \subseteq \mathcal{X}$  with  $|\mathcal{S}| \geq k$  and  $y \in \mathcal{Y}$  such that  $F(a, b) \neq y$  for all  $a, b \in \mathcal{S}$ .

*Proof:* The proof follows from a simple application of Ramsey's theorem for finite sets. First, let's assign some index  $\text{id}(y) \in \{1, \dots, |\mathcal{Y}|\}$  for each  $y \in \mathcal{Y}$ . Then, let

$$\mathcal{H}_b = \{x : x \in \mathcal{X}, \text{id}(F(x, x)) \bmod 3\}$$

and let

$$\mathcal{H}_{\text{maj}} = \arg \max_{\mathcal{H}_b} |\mathcal{H}_b|.$$


I.e.,  $\mathcal{H}_{\text{maj}}$  is the set of elements in  $\mathcal{X}$  that are mapped to  $y$  the most number of times. Then,  $\mathcal{H}_{\text{maj}}$  is a set of size at least  $\frac{|\mathcal{X}|}{3}$ .

We can then think of a graph  $G$  with vertices  $\mathcal{H}_{\text{maj}}$  and an edge between  $(a, b) \in \mathcal{H}_{\text{maj}}$  with the following coloring for edges  $a \neq b$  and  $a < b$ :

$$\text{COL}(a, b) = \text{id}(F(a, b)) \bmod 3.$$

We can then use Ramsey's theorem to find a monochromatic set  $\mathcal{S} \subseteq \mathcal{H}_{\text{maj}}$  with  $|\mathcal{S}| \geq k$  for  $k \geq 0.5 \log_2 \left( \frac{|\mathcal{X}|}{3} \right)$  and color  $c$ . So, what this means is that  $F(a, b) \bmod 3 = c$  for all  $a, b \in \mathcal{S}$  with  $a < b$  for some  $y \in \mathcal{Y}$ . But notice that  $F(a, b) = F(b, a)$  for all  $a, b \in \mathcal{X}$  and so  $F(a, b) = F(b, a) = \text{id}(y) \bmod 3$  for all  $a, b \in \mathcal{S}$  where  $a \neq b$ ! Finally, this means that

$$c = F(a, b) \bmod 3 \quad \text{or} \quad b = F(a, b) \bmod 3$$

as if  $a \neq b$ , the homogeneous edge  $(a, b)$  is colored with  $c$ . If  $a = b$ , then  $F(a, b) \bmod 3 = b$  by definition of  $\mathcal{H}_{\text{maj}}$ ! 

## Strange, isn't it?

At first glance, this lemma may not seem so strange. If we look closer though, we see that this lemma is actually quite surprising! Many every day functions, such as the Diffie-Hellman key exchange, are symmetric functions with an exponentially sized domain, in particular, this means that the lemma applies to them!

So what does this mean for us? I am not entirely sure, but I think it is a good reminder that even the most simple functions can have some very interesting properties!

## Some Maybe Interesting Questions

1. If  $f$  is some polynomial time computable function, can we find a set  $S$  with  $|S| \geq k$  such that  $f(a, b) \neq y$  for all  $a, b \in S$ ? How much time would it take to find such a set? Can we do better than exponential time?
2. As someone interested in cryptography, I wonder if this lemma has any implications for symmetric functions used in cryptography. For example, can we use this lemma to find a set of keys that are not related to each other in some way? Or can does this lemma provide some insight into the security of symmetric functions?