# Non-Adaptive Combinatorial Group Testing

Kirill Levchenko

Consider the following problem. We are given eight coins, of which one is known to be fake. We are also allowed the use of a digital scale. Given that the genuine coins weigh 25 grams, and the fake weighs 20 grams, how many weighings does it take to find the fake?

The solution is three weighings. We start by weighing any four coins. If the weight is not 100 grams, then one of them is fake. Otherwise, the fake is one of the other four. Of the set of four containing a fake, weigh any two. If the weight is not 50 grams, then one of them is fake; otherwise, one of the remaining two is fake. Given two coins of which one is fake, we can find the genuine one with one weighing.

To see that we can do no better, note that each weighing can have only one of two outcomes, depending on whether the fake is among the coins on the scale. Thus, each weighing yields one bit of information. Since we have eight coins, it takes $\lg 8 = 3$ bits to identify any single one.

## Introduction

Abstractly, the problem is to identify a set of "defectives" in a finite universe of items by testing subsets of the universe for defectives, by which we mean testing whether the subset contains a "defective."

There are two orthogonal ways to classify a Group Testing scenario. The first is Combinatorial versus Probabilistic: in Combinatorial Group Testing, the set of defectives is either fixed our bounded by a parameter $k$, while in Probabilistic Group Testing, defectives occur with some probability. The second way is Adaptive versus Non-Adaptive: in Adaptive Group Testing, we specify these tests one at a time, using the outcome of the previous tests, while in Non-Adaptive Group Testing, we must specify all the tests before seeing the outcomes of any of them. The remainder of this note will be concerned with Non-Adaptive Combinatorial Group Testing. Many of the results on this note can be found in [1].

Formally, the *universe* is a finite set $U$ of binary $n$-vectors, or, equivalently, integers 0 to $2^n - 1$. A *test* is a set $P \subseteq U$; we say a test is *positive* with respect to a set $S$ if $P \cap S$ is not empty, and *negative* otherwise. Call a set $\mathcal{P}$ of tests $P_0, \ldots, P_{m-1}$ a *testing scheme*. For a given set $S$, let $Q$ be the set of positive tests with respect to $S$. Call $Q$ the *syndrome* of $S$ with respect to the testing scheme.

We say a testing scheme is *k-separable* if the syndrome of every set of size at most $k$ is distinct. A testing scheme is *weakly k-separable* if the syndrome of every set of size exactly $k$ is distinct. Finally, a testing scheme is *k-disjunct* if for every singleton $\{x\}$ and every set $S$ of size at most $k$ not containing $x$, the syndrome of $\{x\}$ is not contained in the syndrome of $S$.

The value of a $k$-separable scheme is clear, for any set of at most $k$ defectives can be identified using the syndrome, and similarly for a weakly $k$-separable scheme and any set of exactly $k$ defectives. The value of a $k$-disjuct scheme is will be seen shortly.

## Two Simple Bounds

A trivial $k$-separable testing scheme for any $k$ is to test every item individually; that is, let $P_i = \{i\}$

for $i = 0$ to $2^n - 1$. Thus, there exists a $k$-separable testing scheme consisting of $n$ tests, for all $n$ and $k$. Usually, however, we can do much better, as the coin puzzle suggests. Before we consider better testing schemes, it is useful to establish a simple lower bound.

**Proposition 1.** *Any weakly $k$-separable testing scheme requires $kn - k \lg k$ tests.*

*Proof.* There are $\binom{2^n}{k}$ possible sets of defectives. Since each test yields at most one bit of information, the minimum number of tests is

$$\lg \binom{2^n}{k} > kn - k \lg k.$$

□

## A 1-Separable Scheme

When there is exactly one defective, we can achieve the lower bound with the following testing scheme: let $P_i = \{\vec{x} \mid x_i = 1\}$ for $i = 0$ to $n - 1$. It is easy to verify that the syndrome $\vec{t}$ is exactly the defective.

**Proposition 2.** *There exists an (explicitly given) testing scheme for all $n$ and $k = 1$ of size $n$.*

Note also, that a 1-separable testing scheme can be converted to a strongly 1-separable scheme by adding one more test $P_n = U$ to determine if there is a defective at all.

## An Equivalent Definition of $k$-Disjunct

In this section we will give an equivalent definition of $k$-disjunct that will be useful later. We start by establishing some basic facts about syndromes.

**Proposition 3.** *Let a set $S$ have syndrome $Q$, and let a set $S' \subset S$ have syndrome $Q'$. Then $Q' \subset Q$.*

**Proposition 4.** *If $S$ and $S'$ both have a syndrome $Q$, then their union $S \cup S'$ also has syndrome $Q$.*

By Proposition 4, it follows that for any syndrome $Q$, there exists a unique maximal set that yields the syndrome; call such a set *maximally consistent* with $Q$. It is easy to see that for any $S$ with syndrome $Q$, $S$ is a subset of the maximally consistent set of $Q$. We show next that for $|S| \leq k$, equality holds if and only if the testing scheme is $k$-disjunct.

**Proposition 5.** *A testing scheme is $k$-disjunct if and only if for all $S$ of size at most $k$, $S$ is maximally consistent with its syndrome.*

*Proof.* Consider a $k$-disjunct testing scheme. Let $S$ be a set of size at most $k$, and let $M$ be the maximally consistent set of the syndrome of $S$. Toward a contradiction, let $M$ is a proper superset of $S$. Then there is an $x \in M \setminus S$. Furthermore, the syndrome of $\{x\}$ is a subset of $M$, and therefore $S$, contradicting the $k$-disjunct property.

Now consider a testing scheme such that for all $S$ of size at most $k$, $S$ is maximally consistent. Toward a contradiction, let $x$ be an element whose syndrome is a subset of the syndrome of some such $S$, where $x \notin S$. But this implies $S \cup \{x\}$ has the syndrome of $S$, so $S$ is not maximally consistent. □

Proposition 5 gives us an alternate definition of $k$-disjunct.

## Another Definition of $k$-Disjunct

For a set $S$, we say a test $P \in \mathcal{P}$ *isolates* an element $x \in S$ with respect to $S$ if $P \cup S = \{x\}$.

**Proposition 6.** *A testing scheme $\mathcal{P}$ is k-disjunct if and only if for all S of size at most $k+1$ and for every x in S, there is a test $P \in \mathcal{P}$ that isolates x.*

*Proof.* Toward a contradiction, let $\mathcal{P}$ be a $k$-disjunct testing scheme and let $x$ be an element in $S$ of size $k+1$ that is not isolated by some $P$ in $\mathcal{P}$. Then removing $x$ from $S$ does not change the outcome of any test, so $S \setminus \{x\}$ has the same syndrome as $S$. Then by Proposition 3, the syndrome of $\{x\}$ is contained in $S \setminus \{x\}$, contradicting the $k$-disjunct property.

To prove the converse, let $\mathcal{P}$ be a testing scheme with the property that for any sets of size $k+1$ and any element of this set, there is a test in $\mathcal{P}$ that isolates the element with respect to the set. Now assume, toward a contradiction, that $\mathcal{P}$ is not $k$-disjunct, that is, that there is a singleton $\{x\}$ and a set $S$ of size $k$ not containing the singleton, such that the syndrome of $S$ contains the syndrome of $\{x\}$. In particular, for any test $P$ containing $x$, $P \cap S$ must be non-empty, from which it follows that $P$ fails to isolate $x$ with respect to $S \cup \{x\}$. □

## Relation between $k$-Separable and $k$-Disjunct

In this section, we relate $k$-separable and $k$-disjunct.

**Proposition 7.** *$(k+1)$-separable implies k-disjunct.*

*Proof.* Consider a $(k+1)$-separable testing scheme that is not $k$-disjunct. There must be an element $x$ and a set $S$ of size at most $k$ not containing $x$ such that $\{x\}$ and $S$ have the same syndrome. But then $S \cup \{x\}$ has the same syndrome as $S$, violating the $(k+1)$-separable property. □

**Proposition 8.** *k-disjunct implies k-separable.*

*Proof.* Consider a $k$-disjunct testing scheme that is not $k$-separable. There must be distinct sets $S$ and $S'$ of size at most $k$ having the same syndrome. Since they are distinct, there must be an element $x$ is $S$ but not in $S'$. Furthermore, the syndrome of $\{x\}$ is a subset of the syndrome of $S$, and therefore of $S'$, violating the $k$-disjunct property. □

## References

[1] D. Du and F. Hwang. *Combinatorial Group Testing and Its Applications.* Ch. 4 (1993).