隨机演算法 (Probabilistic GT)

Sobel and Groll, Group testing to eliminate efficiently all defectives in a binomial sample, Bell System Tech. J., " (1959) 1179-1252.

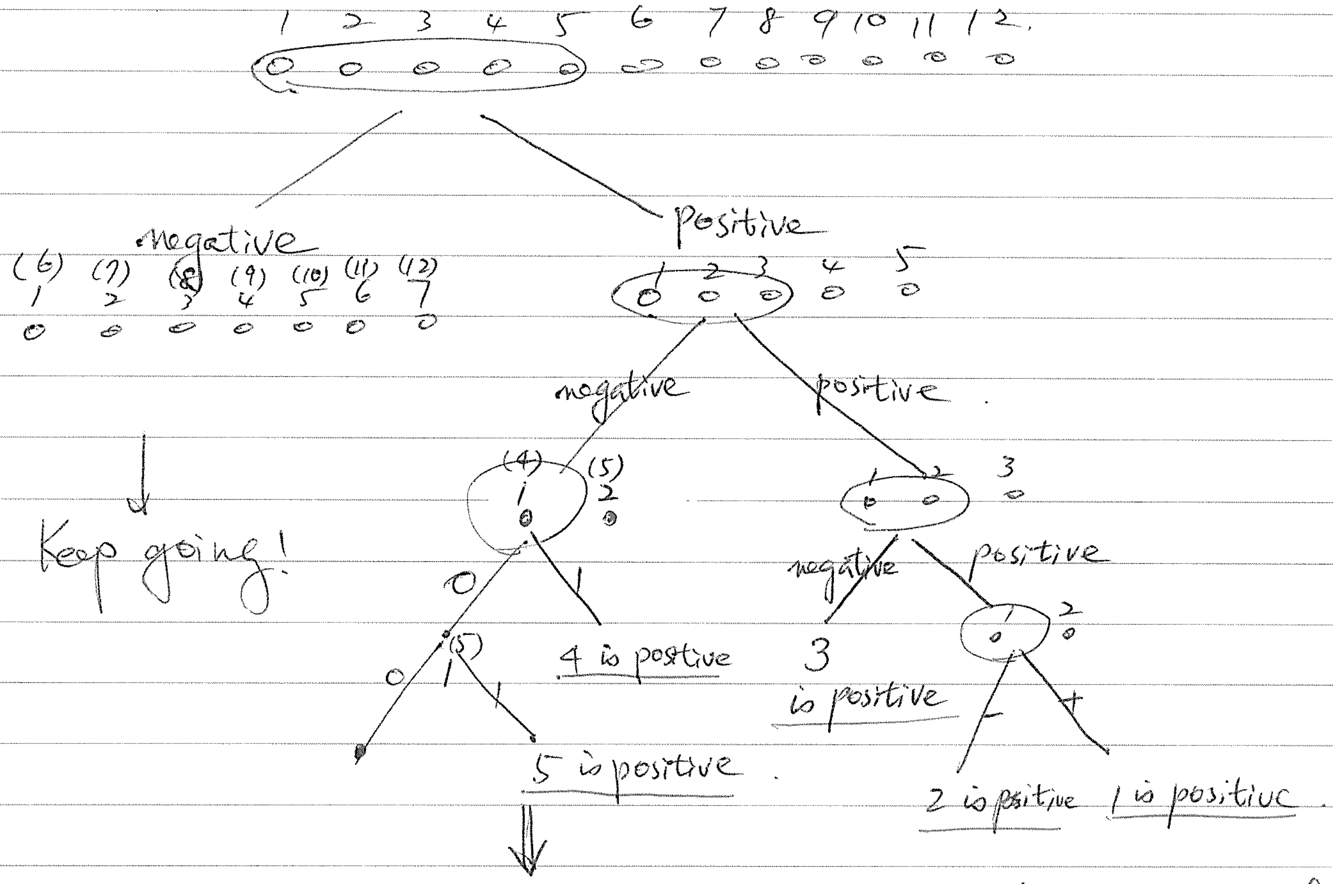
Algorithm

1. There is no restriction on the test group until a group is tested to be positive. Mark the group "Current contaminated group" and denote it by C.
2. The next group $G \subsetneq C$ is to be tested. If G is positive, then replace C (by G) as the current contaminated group. Otherwise the items in G are good and replace C by $C \setminus G$.
3. If "C.C.G." is of size one, identify the item in the group as a defective item. Test any group of unidentified items, if any.

Definition A group testing algorithm is <u>nested</u> if whenever a contaminated group is known, the next group to be tested must be a proper subset of the contaminated group.

Definition (Line algorithm)

A line algorithm always tests a group at the top of the order following an arrangement of unclassified items linearly.



Rearrange the unclassified items linearly

(*) Every nested algorithm can be implemented as a line algorithm, the complexity is about $O(d^2(\log\frac{n}{d})^2)$.

(**) A line algorithm is different from Interval GT algorithm (Start from any interval.)

2-Defective Case $\qquad$ Sep. 27 ~

Fact 1 $\quad M(2,n) \geq \lceil \log_2 \binom{n}{2} \rceil$.

Fact 2 $\quad M(\bar{2},n) \geq \lceil \log_2 \binom{n}{2} + \log_2 n + 1 \rceil$.

) Informatic bound

Definition Let $n_t(d)$ denote the largest n such that the (d,n) problem can be solved in t tests.

(*) Determining $n_t(2)$ is very hard, but $n_t(1) = 2^t$ which is easy.

Definition For $t \geq 1$, let α_t denote the integer such that

$$\binom{\alpha_t}{2} < 2^t < \binom{\alpha_t + 1}{2}. \quad \text{(Equal sign is not possible!)}$$

e.g. $t=3$, $\alpha_t = 4$; $t = 5$, $\alpha_t = 8$.

Fact 3. $\quad \alpha_t = \lfloor 2^{\frac{t+1}{2}} - \frac{1}{2} \rfloor + 1$.

Proof. $\quad \alpha_t < 2^{\frac{t+1}{2}} + \frac{1}{2} < \alpha_t + 1$

$$\Rightarrow \binom{\alpha_t}{2} < \frac{(2^{\frac{t+1}{2}} + \frac{1}{2})(2^{\frac{t+1}{2}} - \frac{1}{2})}{2} < \binom{\alpha_t + 1}{2}$$

$$\frac{(2^{\frac{t+1}{2}} + \frac{1}{2})(2^{\frac{t+1}{2}} - \frac{1}{2})}{2} = 2^t - \frac{1}{8}$$

∎

Fact 4. $\quad \binom{\alpha_{t+1}}{2} - \binom{\alpha_t - 1}{2} > 2^t$.

Proof. Direct checking by using t is even or odd. ... ∎

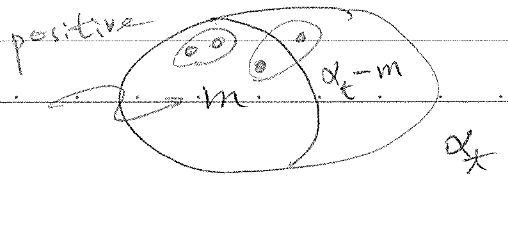
<u>Theorem 5.</u> $n_t(2) \le \alpha_t - 1$ for $t \ge 4$.

<u>Proof.</u> By induction on t. First, $n_4(2) \le 5 = \alpha_4 - 1$. Let T be an arbitrary algorithm. We prove $M_T(2, \alpha_t) > t$ and thus $n_t(2) \le \alpha_t - 1$.

Now, find a set of m items and test it. First, if $m < \alpha_t - (\alpha_{t-1} - 1)$ and the outcome is negative, then the problem is reduced to a $(2, \alpha_t - m)$ problem (after one test). Since $\alpha_t - m > \alpha_{t-1} - 1$, at least t more tests are needed (by induction ($n_{t-1}(2) \le \alpha_{t-1} - 1$) to find two positives. [On the other hand, if the outcome is positive, then we may need less tests. But, the negative outcome provides the worse case.]

Second, if $m \ge \alpha_t - (\alpha_{t-1} - 1)$, consider the <u>positive outcome</u>. The set of samples (2-subsets positive) has size (after 1st test)

$$\binom{\alpha_t}{2} - \binom{\alpha_t - m}{2} \ge \binom{\alpha_t}{2} - \binom{\alpha_{t-1} - 1}{2} > 2^{t-1} \text{ (by Fact 4).}$$



Hence, t more tests are needed. This implies $M_T(2, \alpha_t) > t$. ∎

Some more progress has been made which can be found in literature. But, solving $(2,n)$ problem is still far from being done.

3- Defective Case

Theorem 6. There exists an algorithm T such that the $(3, h_t)$ problem can be solved in t tests where

$$h_t = \begin{cases} t+1 & \text{for } t \le 7, \\ h_{3k+1} + 2^{k-1} & \text{for } t = 3k+2 \ge 8, \\ n_{2k}(2) + 2^{k-2} + 1 & \text{for } t = 3k \ge 9, \text{ and} \\ \left\lfloor (n_{2k+1}(2) + n_{2k}(2))/2 \right\rfloor + 3 \cdot 2^{k-2} + 1 & \text{for } t = 3k+1 \ge 10, \end{cases}$$

except $h_{12} = 26$, $h_{13} = 32$, $h_{14} = 40$ and $h_{15} = 52$.

As you can see, finding $n_t(3)$ is getting more complicate. If d is getting larger comparing to n, then we may as well use the individual test. The following result shows that group testing is still working well when $n > 3d$.

<u>Theorem 2</u> $M(d,n) < n-1$ for $n > 3d$.

<u>Proof.</u> It suffices to prove $M(d,3d+1) < 3d$. (If $n > 3d+1$, let $n-(3d+1) = d'$. Now, take individual test on these d' items. It takes $d' + M(d,3d+1)$ tests to find all positive items. Hence, the total number of tests will be <u>less</u> <u>than</u> $d'+3d = n-1$.)

The proof of $M(d,3d+1) < 3d$ follows by showing $3d-1$ tests will be enough to find all the positive items. Now, we take two items (a group) (using line algorithm) and test them. Clearly, it takes one test to find two negative items and two tests to find a positive item. So, if we do need $3d$ tests to find all the positives, then all items are identified. But, in fact, the last test can be waived. Now, suppose that we use $3d-1$ tests already. Then, either d positive items have been identified or $c < d$ positives are identified. Note that $c = d-1$. (Otherwise, say $c = d-2$, then $[(3d-1)-(d-2)]/2 + 2(d-2) \leq 3d-1$.) In the first case, the other items are negative, no further test is necessary. For the second case, we have identified the negative items using $(3d-1)-2(d-1)$ Tests which is "$d+1$ tests". Therefore, the total items are $2(d+1)+(d-1)=3d+1$.

e.g. $M(5,16) < 15$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16

①	1, 2	Good.	
②	3, 4	Defective	
③	3	Defective. ✓	
④	4, 5	Good	
⑤	6, 7	Defective.	
⑥	6	Defective ✓	
⑦	7, 8	Good	
⑧	9, 10	Defective	
⑨	9	Good	⇒ 10 Defective ✓
⑩	11, 12	Good	
⑪	13, 14	Defective	
⑫	13	Defective ✓	
⑬	14, 15	Defective	
⑭	14	Defective ✓	
	15, 16	Good	(No test!)

13, 14, 15, 16

13	Defective.
14, 15	Good.
16	Defective

13, 14, 15, 16

⑪	13, 14 .	Defective.
⑫	13 .	Defective. ✓
⑬	14, 15	Defective
⑭	14 .	Good ⇒ 15 . defective. ⇒ 16 Good

Another example

$M(4,13)$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13

①	1, 2 :	$-$	
②	3, 4 :	$+$	
③	3	$-$	$\Rightarrow$ 4 ✓ $+$
④	5, 6	$+$	
⑤	5	$+$	✓
⑥	6, 7	$+$	
⑦	6	$-$	$\Rightarrow$ 7 ✓ $+$
⑧	8, 9	$-$	
⑨	10, 11	$-$	
⑩	12, 13	$+$	
⑪	12	$-$	$\Rightarrow$ 13 $+$

Theorem 8 (The best result so far)

$M(d,n) = n-1$ for $n \le \frac{21}{8}d$, i.e, if $n > \frac{21}{8}d$, then $M(d,n) < n-1$.

Ref. Ding Zhu Du and F.K. Hwang, SIAM J. AlG. DISC. METH., Vol. 3, No. 4, Dec. 1982, 523-528.