(1)

3,9

~3,10

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

Let $f(m,n)$ denote the number nested algorithms when the group testing consists n items and a contaiminated group is of size m.

$$f(0,0) = 1$$

$$f(0,n) = \sum_{k=1}^{n} f(0, n-k) \cdot f(k, n)$$

$$f(1,n) = f(0, n-1)$$

$$f(m,n) = \sum_{k=1}^{m-1} f(m-k, n-k) \cdot f(k, n).$$

(k is the size of the group to be tested.)

$$C_k = \frac{1}{k}\binom{2k-2}{k-1} \text{ — Catalan}$$

#

$$C_k = \sum_{i=1}^{k-1} c_i C_{k-i} \quad \text{for } k \geq 2.$$

Theorem

$$f(0,n) = C_{n+1} \prod_{i=1}^{n-1} f(0,i) \quad \text{for } n \geq 2;$$

$$f(m,n) = C_m \prod_{i=1}^{m} f(0,n-i) \quad \text{for } 1 \leq m \leq n.$$

Proof.

Let $F(n) = f(0,n)$. Then

$$F(n) = 4^{2^n-1} \prod_{i=1}^{n} \left\{ 1 - \frac{3}{2(i+1)} \right\}^{2^{n-i}}$$

$$\approx \frac{1}{4} \alpha^{2^n} \left\{ 1 + \frac{3/2}{n} + o\left(\frac{1}{n}\right) \right\}$$

where $\alpha = 1.526753\cdots$

Also, $F(n) > \frac{1}{4} \alpha^{2^n}$.

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

Proof: It suffices to prove that $M(d, 3d+1) < 3d$.

(If $n > 3d+1$, let $n = 3d+1+i$, $i > 0$)

Test the i items individually and add the number of tests for $M(d, 3d+1)$, we have

$$i + M(d, 3d+1) < i + 3d = n - (3d+1) + 3d = n-1.$$

Algorithm

Let T be a nested (line) algorithm which always tests a group of two items. Since T identifies a defective by using at most two tests and identifies good items in pairs except the last item (when $3d+1$ is odd) or those good items appearing after the last defective are identified by deduction with testing. Hence T takes at most $3d$ tests. Assume T takes $3d$ tests. This must be the case when we already use $3d-2$ tests and

Now, only one test is needed. Hence, we have the proof.

最早猜測 $n \le 3d$ 就需要用 個別測。

$\boxed{n \le 3d \text{ Conjecture (Hu, Hwang and Wang)}}$

A boundary problem for group testing, SIAM. J. Alg. Discrete Methods 2 (1981) 81–87.

Theorem

$M(d,n) = n-1$ provided $d < n \le \lfloor \frac{5d+1}{2} \rfloor$.

Theorem (Improved) (Du and Hwang)

$M(d,n) = n-1$ provided $n \le \frac{21}{8} d$.

Minimizing a combinatorial function, SIAM J. Alg. Discrete Methods 3 (1982) 523–528.

Theorem (?) You?

3, 18 ①

Notations

For an algorithm α, let N_α("s") denote the number of tests used by the algorithm α on the sample "s".

Let $S(d,n)$ denote the set of samples of n items containing d defectives and $M_\alpha(d|n) = \max_{s \in S(d,n)} N_\alpha(s)$.

An algorithm α is called a c-competitive algorithm if $\exists$ a constant a such that for $0 \le d < n$, $M_\alpha(d|n) \le c \cdot M(d,n) + a$.

competitive ratio of α

(*) 在 $d = n$ 時 $M(n,n) = 0$ but $M_\alpha(n|n) \ge n$ for any algorithm α.

(**) $M(d|n)$ 是 Competitive Model 的最佳解。

Bisecting (Algorithm A1) $|S| = n$

(*) The principle of the bisecting algorithm is that at each step, if a contaminated subset X of S is discovered, then bisect the set X and test the resulting two subsets X' and X''.

(*) It is better off to let $|X'| = 2^{\lceil \log_2 |X| \rceil - 1}$ and $X'' = X \setminus X'$.

For example, if $|X| = 100$, let $|X'| = 2^{\lceil \log_2 100 \rceil - 1} = 64$

and $|X''| = 36$.

G : set of good items
Q : Queries (Queue)
D : defectives

Algorithm A1 :

input S ;
$G \leftarrow \phi$;
$D \leftarrow \phi$;
if S is pure
  then $G \leftarrow S$ and $Q \leftarrow \phi$
  else $Q \leftarrow \{S\}$
repeat
  pop the frontier element X of queue Q;
  "bisect" X into X' and X'';

if X is contaminated then Test(X');
{ if X' is pure, then X'' is contaminated }
for $Y \leftarrow X'$ and (or) X'' do begin
    if Y is pure, then $G \leftarrow G \cup Y$;
    if Y is contaminated and $|Y|=1$
        then $D \leftarrow D \cup \{Y\}$;
    if Y is contaminated and $|Y|>1$
        then push Y into Q
  end-for;
until $Q = \emptyset$
end-algorithm.

(Fact) $M_{A_1}(d \mid n) \le 2n-1$ for any d.

令 T^* 为 T (BFS) tree 的 Computation Tree.

因为 n 个 leaves (判断全部的可能性),

所以 Worst case 是 $2n-1$.

(***) If n is a power of 2, then for $1 \le d \le n$,

$$M_{A_1}(d|n) \le 2d(\log \frac{n}{d} + 1) - 1.$$

Proof.

If $\boxed{n \text{ is not a power of } 2}$, then

$$M_{A_1}(d|n) \le 2d(\log \frac{n}{d} + 1) \underline{+1}.$$

(顶多加两个 tests)

$\Downarrow$

a

$$M_{A_1}(d|n) \le 2M(d,n) \underline{\underline{+5}}.$$

A_1 is a 2-competitive algorithm.

Doubling algorithm $\boxed{A_2}$

Strategy

It tests disjoint sets of sizes $1, 2, \cdots, 2^i$ until a contaminated set is found.

Doubling

The algorithm identifies

$1+2+\cdots+2^{i-1}=2^i-1$ good items and

a contaminated set with $i+1$ tests by a binary search. And then

indentify a defective in i tests.

(☆) In total, the algorithm spends

$2i+1$ tests and identifies 2^i items.

$(i=1,2,\cdots)$.

DIG

⋮

Theorem For $1 \le d \le n-1$, $M_{A_2}(d|n)$

$\le 2M(d,n)+4$.

Jumping (A_3)

Doubling

$$1, 2, 2^2, \cdots, 2^i$$

Jumping

$1+2, 4+8, \cdots, 2^i + 2^{i+1}$ for even i, i.e.,

$1+2, 2^2+2^3, 2^4+2^5, \cdots, 2^{2t}+2^{2t+1}$ for $t \in \mathbb{N}$.

identifies $2^i - 1$ good items with $\frac{i}{2}$ tests (instead of i)

and finds a contaminated set of size

$3 \cdot 2^i$ (instead of 2^i).

$2^i \quad 2^{i+1}$ 3-5 test

Theorem $M_{A_3}(d|n) < 1.65\, M(d,n) + 10.$