Finding A Divisible Pair

Stelian Ciurea[1] Erik D. Demaine[2] Corina E. Pătraşcu[3]
Mihai Pătraşcu[2]

[1]University of Sibiu, Romania

[2]MIT CSAIL

[3]Department of Mathematics
Harvard University

presented at FUN with Algorithms 2004

The Divisible Pair Problem

Theorem

For any $S \subset \{1, \dots, 2n\}$*, with* $|S| = n + 1$*, there exist* $a, b \in S$ *such that a divides b.*

Proof.

$\forall q \in \{1, 3, \dots, 2n-1\}$, let $B_q = \{q \cdot 2^i \mid 0 \le i \le \log_2 \lfloor n/q \rfloor\}$.
There are n such B_q's; they are all disjoint and their union is $\{1, \dots, 2n\} \Rightarrow$ by the pigeon hole principle, since $|S| = n + 1$, at least two numbers must be in the same set. $\square$

Note. So, not only do we have a divisible pair, but one of the numbers is a power of 2 multiple of the other.

Algorithmic Challenge

Problem

Find the divisible pair using membership queries – the only allowed questions are whether a number is in the set or not.

Naive Upper Bound: $\frac{3}{2}n + O(1)$

Probe the i's in increasing order:

- if $i > n$ is odd, don't query it because from the proof, such an i can't be part of a divisible pair;
- for all other i's, test whether $i \in S$; if so, test membership of all multiples of the form $2^j i$.

Claim: Each i is tested once because if we assume that an i is examined twice $\Rightarrow \exists x, y \in S$ such that $i = x \cdot 2^a = y \cdot 2^b$. But then $y = x \cdot 2^{a-b}$ and when $i = x, j = a - b$, the procedure must have stopped and returned the pair (x, y).

Lower Bound: $\frac{4}{3}n - O(1)$

The adversary has the following strategy, which will force any player to make at least $\frac{4}{3}n - O(1)$ queries.

Strategy

- $\forall x \in \{1, \ldots, \frac{2}{3}n\}$ say NO (if there still are NO's left) because they have too many multiples in $\{1, \ldots, 2n\}$ and they are difficlut to control;
- $\forall x \in A = \{\frac{2}{3}n + 1, \ldots, n\}$, for pairs of the form $(x, 2x)$ say YES to the first asked to force the player to ask about the second, and say NO for the second as long as there are still NO's left;
- $\forall x \in \{n, \ldots 2n\} \cap \mathrm{complement}(2A)$ say YES because they have no multiples.

Analysis of Lower Bound

- this adversary will not generate a divisible pair before exhausting all $n-1$ NO's (since there are $n-1$ numbers outside S);
- we get a negative answer only from one of the first 2 stages of the adversary's strategy;
- we get at most $\frac{2}{3}n$ NO's from the first stage
- $\Rightarrow$ we get at least $n-1-\frac{2}{3}n=\frac{n}{3}-1$ NO's from the second stage;
- since for every NO in the second part we also get a YES, then we get at least $\frac{n}{3}-1$ YES's until we stop.

Therefore, we have to make at least $n-1+\frac{n}{3}-1=\frac{4}{3}n-O(1)$ queries before we stop.

Upper Bound: $\left(\frac{4}{3} + \frac{1}{24}\right) n + O(1)$

Algorithm

Probe the i's in increasing order.

- **Phase 1**
  for $1 \leq i \leq \frac{2}{3}n$, if $i \in S$, probe all multiples
- **Phase 2**
  for $\frac{2}{3}n < i \leq n$:
  - if $2i$ already probed, then ignore i because:
    - either $2i \in S$, in which case, we already stopped;
    - or $2i \notin S$, in which case, it doesn't matter whether i is in S or not, because $2i$ is its only multiple and so i can't be part of a divisible pair;
  - else, probe i; if $i \in S$, probe its multiple (i.e. $2i$).

Upper Bound Analysis

- if the algorithm stops in Phase 1:
  - we got at most $n-1$ NO's so far;
  - at most one query for $x > \frac{2}{3}n$ can be YES, because since we stop in Phase 1, the only way to probe an $x > \frac{2}{3}n$ is if it is a multiple of something $< \frac{2}{3}n$, in which case we found our divisible pair;
  - using the original theorem, if we probe $\frac{n}{3}+1$ numbers from the set $\{1, \dots, \frac{2}{3}n\}$, we have to find a divisible pair; so, at most $\frac{n}{3}+1$ of the queries from this phase can be answered YES before we stop.

Therefore, we will make at most $\frac{4}{3}n + O(1)$ total queries.

Upper Bound Analysis continued

- if the algorithm stops in Phase 2:
  - all $i \leq \frac{2}{3}n$ were probed;
  - for $i > \frac{2}{3}n$, for every NO, there exists at most a YES because:
    - if i and $2i$ are in S, then we stop, so we have 2 YES's;
    - if $i \notin S$ or $2i \notin S$ we simply ignore the other because it can't be part of a divisible pair, so it doesn't matter whether it gives a YES or a NO $\Rightarrow$ at most one YES;

Upper Bound Analysis – Phase 2 continued

Notation. Let $Y \subset \{1, \ldots \frac{2}{3}n\}$, $Y =$ numbers in S discovered in Phase 1 and let $N \subset \{1, \ldots 2n\}$, $N =$ numbers probed in Phase 1 which are not in S $\Rightarrow$ there are $n - 1 - |N|$ numbers outside S which are not yet discovered.

So, since there are at most as many YES's as there are NO's and when we stop we get two YES's for the divisible pair, then the total number of queries during Phase 2 is at most $2 + 2(n - 1 - |N|) = 2n - 2|N|$.

Therefore, the total number of queries is at most $|Y| + |N| + 2n - 2|N| = 2n - (|N| - |Y|)$.

So, it is enough to prove that $|N| - |Y| \geq \left(\frac{2}{3} - \frac{1}{24}\right) n$.

Upper Bound Analysis – Phase 2 continued

Lemma

$|N| - |Y| \geq \left(\frac{2}{3} - \frac{1}{24}\right) n.$

Proof

By our definition, it is clear that N contains exactly $\frac{2}{3}n - |Y|$ numbers from $\{1, \dots, \frac{2}{3}n\}$. Let $M = N \cap \{\frac{2}{3}n + 1, \dots, 2n\}$. Then, our bound is equivalent to $|M| \geq 2|Y| - \frac{n}{24}$.

- M contains all multiples greater than $\frac{2}{3}n$ of numbers from Y, in particular all power of 2 multiples;
- in M, there are at least 2 power of 2 multiples for every number in $Y \cap \{1, \dots, \frac{n}{2}\}$ and at least 1 multiple for the rest of the numbers in Y;
- however, for numbers in $Y \cap \{\frac{n}{2} + 1, \dots, \frac{2}{3}n\}$, M must also contain their triple;

Upper Bound Analysis – Phase 2 continued

Proof (continued)

- thus, for every number in Y, at least two of its multiples are in M;
- no two numbers from Y can have the same power of two multiple so the only double counting can arise when $3x = 2^i y$, for some $y \in Y$ and $x \in Y \cap \{\frac{n}{2}+1, \ldots, \frac{2}{3}n\}$;
- $3x \in \{\frac{3}{2}n+3, \ldots, 2n\}$, so $2^{i-1}y \in \{\frac{3}{4}n+1, \ldots, n\}$ and it must also be a multiple of 3;
- also, $2^{i-1}y$ must be even because $y \le \frac{2}{3}n$ so $i \ge 2 \Rightarrow$ $2^{i-1}y$ must be a multiple of 6 in the range $\{\frac{3}{4}n+1, \ldots, n\}$;
- since there are $\frac{n}{24}$ such possibilities and each one defines at most one double-counted multiple $\Rightarrow |M| \ge 2|Y| - \frac{n}{24}$ which completes our proof.