

Notes: TODO - The 3SUM problem

Carolin Zöbelein*

August 2, 2020

DRAFT

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Keywords: XXX, YYY, ZZZ

2010 Mathematics Classification: XXX, YYY, ZZZ

Preamble

The following content is a sketch for discussion purposes only, without warranty for mathematical completeness.

1 Introduction

We consider the 3SUM problem and its limited to only integers version **Int3SUM**, a variant of this one of the most basic problems in algorithm design, which are defined as follows:

Definition 1 (3SUM). *Given a set $S \subset \mathbb{R}$, determine if there exists $a, b, c \in S$ such that $a + b + c = 0$.*

Definition 2 (Int3SUM). *Given a set $S \subseteq \{-U, \dots, U\} \subset \mathbb{Z}$, determine if there exists $a, b, c \in S$ such that $a + b + c = 0$.*

Since numerous problems can be reduced from 3SUM [1] [2], it has raised wide interest. Such problems which are at least as hard as 3SUM are called *3SUM-Hardness*. Some of them are ...

- Given a n -point set in the plane, determine if it contains three collinear points. It was shown by Gajentaan and Overmars [1] that it can be solved in $\mathcal{O}(n^2)$ time.
- Given a set of n triangles in the plane, determine if their union contains a hole or not, respectively compute the area of their union which can be also solved in $\mathcal{O}(n^2)$ time [1].

*The author believes in the importance of the independence of research and is funded by the public community. If you also believe in this values, you can find ways for supporting the author's work here: <https://research.carolin-zoebelein.de/funding.html>, Email: contact@carolin-zoebelein.de, PGP: D4A7 35E8 D47F 801F 2CF6 2BA7 927A FD3C DE47 E13B, <https://research.carolin-zoebelein.de>

- Given two n -point sets $X, Y \subset \mathbb{R}$, each of size n , determine if all elements in $X + Y = \{x + y \mid x \in X, y \in Y\}$ are distinct. It can be solved in $\mathcal{O}(n^2 \log(n))$ time [2]. This problem, including its stronger version with sorting $X + Y$, is used for the conditional lower bounds of problems by Barequet and Har-Peled in [2] and are classified as *Sorting $X + Y$ -Hard*. It's an open question if it can be solve in $o(n^2 \log(n))$ [3] [4] [5].
- Given two n -edge convex polygons, determine if one can be placed inside the other by translation and rotation which can also be solved in $\mathcal{O}(n^2 \log(n))$ time [2].
-

1.1 Related Work

The lower bound of computational complexity of 3SUM is a long standing unsolved question. Over the years, work has been done and several results were achieved.

- A trivial hash table algorithm can solve 3SUM in $\mathcal{O}(n^2)$ time.
All elements $s \in S$ are hashed into a table T_h by $h(s)$. Then we go for all $(h(s_i) + h(s_j))$ through the hash table and check for $-(h(s_i) + h(s_j))$.
- Erickson [6] credited Seidel [7] that **Int3SUM** can be solved with fast Fourier transform in $\mathcal{O}(n + U \log(U))$ time, for a universe size U (an integer range $[-U, \dots, U]$).
- Baran, Demaine and Pătraşcu [8] showed that **Int3SUM** can be solve in $\mathcal{O}\left(n^2 / (\log(n) / \log \log(n))^2\right)$ time with high probability on the word RAM with $U = 2^\omega$ and $\omega > \log(n)$ is the machine word size. It uses a mixture of randomized universe reduction by hashing, word packing and table lookups.
- Erickson [6] and Ailon and Chazelle [9] proved that any 3-linear decision tree for 3SUM has depth $\Omega(n^2)$ and matching $\Omega(n^{\lceil k/2 \rceil})$ lower bounds in some particular models.
- Grønlund and Pettie [10] showed that 3SUM can be solved in subquadratic time by a deterministic algorithm which runs in $\mathcal{O}\left(n^2 (\log(n) / \log \log(n))^{2/3}\right)$ time and a randomized algorithm which runs in $\mathcal{O}\left(n^2 (\log \log(n))^2 / \log(n)\right)$ expected time with high probability. They also showed that it exists a 4-linear decision tree with deph $\mathcal{O}\left(n^{3/2} \sqrt{\log(n)}\right)$.
- Freund [11] developed a 3SUM algorithm by eliminating one of Grønlund and Pettie's conceptual building blocks and lead to a bound for running time of $\mathcal{O}(n^2 \log \log(n) / \log(n))$.
- Gold and Sharir [12] showed that the randomized 4-linear decision tree complexity of 3SUM is $\mathcal{O}(n^{3/2})$.
- Chan [13] gave an algorithm that solved 3SUM in $\mathcal{O}\left((n^2 / \log(n)) (\log \log(n))^{\mathcal{O}(1)}\right)$ time and obtained subquadratic results on some 3SUM-hard problems.
- Kane, Lovett, and Moran [14] constructed a linear decision tree that solves the 3SUM problem with $\mathcal{O}(n \log^2(n))$ queries which compare the sums of two 3-subsets.
- Kopelowitz, Pettie and Porat [15] conjectured that 3SUM is unsolvable in $\mathcal{O}(n^{2-\omega(1)})$ expected time.

1.2 Our Results

TODO

1.3 An Overview

TODO

2 TODO: Old text

2.1 Introduction

In this notes we will give a sketchy discussion of an algorithm for solving the 3SUM problem.

2.2 Preliminaries

Some basic definitions and preparatory work.

2.2.1 Problem Statement

Let L be a list of n integers n_i , $i \in \mathbb{N} : i = 1, \dots, n$. All n integers are given in a binary representation of b bits each. Negative integers are given by their binary two's complement representation.

Definition 3 (3SUM Problem). *The 3SUM problem asks, if a given list L of n integers contains three elements n_i , n_j and n_k , such that $n_i + n_j + n_k = 0$ for $i \neq j \neq k$.*

2.2.2 Binary addition

Our algorithm will work on the binary representation level of the given integers n . For this, we will have a short look at the different possible cases of bits for the addition of three numbers n_i , n_j and n_k .

Notations. We will index the bits b_i of n_i from right to left with $q = 0, \dots, b - 1$ and denote the q 'th bit of n_i by $b_i[q]$.

Zero sum binary addition

Since we are interested in the special case of zero sum of three numbers, we discuss this case, at first.

We start with bit $q = 0$. Here, we always have exactly three bits for our addition (the 0'th bit of each of our numbers n_i , n_j and n_k). Since, we are only interested in the zero sum case, we have to think for which combinations of three bits we can get 0, at all. This are:

1. $\mathcal{C}_{0,0} = 3, \mathcal{C}_{0,1} = 0, c_0 = 0$: $(0, 0, 0)$
2. $\mathcal{C}_{0,0} = 1, \mathcal{C}_{0,1} = 2, c_0 = 1$: $(1, 1, 0), (1, 0, 1), (0, 1, 1)$

$\mathcal{C}_{q,0}$ denotes the number of 0's and $\mathcal{C}_{q,1}$ the number of 1's in our three q bits $b_i[q]$, $b_j[q]$ and $b_k[q]$. c_q denotes the carrier of the given bit addition $b_i[q] + b_j[q] + b_k[q]$.

If we have $c_0 = 0$, then, for bit $q = 1$, we will have the same situation like for bit $q = 0$. Hence, let's look at the case of having a carrier $c_0 = 1$ and what will happen then for bit $q = 1$. Because of the carrier $c_0 = 1$, we have one bit more, hence four bits, now. We get zero sum for:

1. $\mathcal{C}_{1,0} = 0, \mathcal{C}_{1,1} = 4, c_1 = 10$: $(1, 1, 1, 1)$
2. $\mathcal{C}_{1,0} = 2, \mathcal{C}_{1,1} = 2, c_1 = 1$: $(1, 1, 0, 0), (1, 0, 1, 0), (1, 0, 0, 1)$

In the first case with carrier $c_1 = 10$, we will get for $q = 2$ the three bit case again, and for $q = 3$ the four bit case again. Depending if we choose $(0, 0, 0)$ or $\{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$ for bit $q = 2$, we will finally get for bit $q = 3$ a four respectively five bit case (see also table 1). In the second case with carrier $c_1 = 1$, we will get for $q = 2$ a four bit case again.

So, let's assume we have a five bit case for bit $q = 3$ and get zero sum for:

1. $\mathcal{C}_{3,0} = 1, \mathcal{C}_{3,1} = 4, c_3 = 10$: $(1, 1, 0, 1, 1), (1, 1, 1, 0, 1), (1, 1, 1, 1, 0)$

Table 1: Binary addition of three numbers with maximal carrier.

		1		1				
1	0	1	0	1	0	1		
...	...	1	1	1	1	1	1	
...	...	1	1	1	1	1	1	
...	...	0	0	0	0	1	0	
...	...	0	0	0	0	0	0	
...	...	5	4	3	2	1	0	q

2. $\mathcal{C}_{3,0} = 3, \mathcal{C}_{3,1} = 2, c_3 = 1$: $(1, 1, 0, 0, 0)$

Here, we have a very similiar situation like for the case with four bits at $q = 1$. If we have carrier $c_3 = 10$ for bit $q = 3$, we will have the three bit case for $q = 4$ and the four or five bit case for $q = 5$. If we have carrier $c_3 = 1$, we will have the four bit case for bit $q = 4$.

After this short discussion of possible cases, we will write this in a more formal way, next.

Zero sum bit rules

For the binary addition of the three numbers n_i, n_j and n_k we can put the results for the zero sum case together to a complete set of rules. We will write short $\mathcal{C}_{q,b}^{c_q} := (\mathcal{C}_{q,0}, \mathcal{C}_{q,1}, c_q, b)$ with b be the number of bits of the belonging case, for our possible bit cases. With this we can write:

- Case: $b = 3$

$$\begin{aligned} \mathcal{C}_{q,3}^0 &\rightarrow \{\mathcal{C}_{q+1,3}^0, \mathcal{C}_{q+1,3}^1\} \\ \mathcal{C}_{q,3}^1 &\rightarrow \{\mathcal{C}_{q+1,4}^{10}, \mathcal{C}_{q+1,4}^1\} \end{aligned} \tag{1}$$

- Case: $b = 4$

$$\begin{aligned} \mathcal{C}_{q,4}^{10} &\rightarrow \{\mathcal{C}_{q+1,3}^0 \rightarrow \{\mathcal{C}_{q+2,4}^{10}, \mathcal{C}_{q+2,4}^1\}, \mathcal{C}_{q+1,3}^1 \rightarrow \{\mathcal{C}_{q+2,5}^{10}, \mathcal{C}_{q+2,5}^1\}\} \\ \mathcal{C}_{q,4}^1 &\rightarrow \{\mathcal{C}_{q+1,4}^{10}, \mathcal{C}_{q+1,4}^1\} \end{aligned} \tag{2}$$

- Case: $b = 5$

$$\begin{aligned} \mathcal{C}_{q,5}^{10} &\rightarrow \{\mathcal{C}_{q+1,3}^0 \rightarrow \{\mathcal{C}_{q+2,4}^{10}, \mathcal{C}_{q+2,4}^1\}, \mathcal{C}_{q+1,3}^1 \rightarrow \{\mathcal{C}_{q+2,5}^{10}, \mathcal{C}_{q+2,5}^1\}\} \\ \mathcal{C}_{q,5}^1 &\rightarrow \{\mathcal{C}_{q+1,4}^{10}, \mathcal{C}_{q+1,4}^1\} \end{aligned} \tag{3}$$

Notations. We write $\mathcal{C}_{q,b}^{c_q}(0)$ to get the value of $\mathcal{C}_{q,0}$ from $\mathcal{C}_{q,b}^{c_q}$, $\mathcal{C}_{q,b}^{c_q}(1)$ to get the value of $\mathcal{C}_{q,1}$ from $\mathcal{C}_{q,b}^{c_q}$, $\mathcal{C}_{q,b}^{c_q}(c)$ to get the value of c_q from $\mathcal{C}_{q,b}^{c_q}$ and $\mathcal{C}_{q,b}^{c_q}(b)$ to get the value of b from $\mathcal{C}_{q,b}^{c_q}$.

2.3 The 3SUM Algorithm

Now we want to describe the final algorithm.

2.3.1 Idea

We want to outline our 3SUM algorithm, at first.

Basic concept

Typical 3SUM algorithms like XXX [?] and YYY [?] works on the decimal base angel of view by going in

different ways through the numbers n of the given list L . In our algorithm, we move away from this approach. Instead, we will go through the binary representation of the given numbers. So, the final algorithm will be mainly dominated by the number of bits b of integer representation which we have to examine and not by the number of elements n of L .

What we do. Since, we examine the 3SUM problem, we consider the binary addition of three integer numbers. We know from our previous discussion that only certain combinations of bits resulting in zero sums and that the possibilities for bits of position q depend from the possibilities for bits of position $q - 1$. We begin with the whole list L and split it into two sublists. One which contains all numbers with 0 at bit position $q = 0$ and the other list which contains all numbers with 1 at bit position $q = 0$. We know, that for bit $q = 0$ we have two possible rules which lead to a zero bit ($\mathcal{C}_{q,3}^0, \mathcal{C}_{q,3}^1$). We check for the first rule if the two sublists have enough elements (we also save the amount of numbers of each sublist) to satisfy this rule. If yes, then we save this rule, if no, we throw it away. Then we do the same for the second rule. So, we have solved the 3SUM problem for bit $q = 0$.

In the next step, we split again each of our sublists into two further sublists. At this time now, we do it for bit $q = 1$ instead of bit $q = 0$, of course. Then, we go through the list of our saved rules. We take the first saved rule and determine its two successor rules. We take the first successor rule and check our new sublists if they have enough elements to satisfy this rule. If yes, we also have to check if the necessary amounts of numbers in this sublists are not in conflict with the amounts from our step before. If no, then save this new rule, else throw it away. In the same way, we do it for all possible rule successors.

After repeating this steps for all bits b of our bit representation, we get left with the set of all rules and their belonging numbers which solve the 3SUM problem.

Hence, this algorithm depends on the number of bits of bit representation but not primary on the number of elements of n . Each time if we do a list split we can already count the number of elements of this list during splitting and no search, at no point in the algorithm, is necessary, if we implement it with two binary trees (one for the number lists and one for the rules) and pointers which always point to the last necessary elements of this trees.

Algorithm

Pseudocode

Complexity

Acknowledgement

Thanks to the private donators who financially support this work.

License

<https://creativecommons.org/licenses/by-sa/4.0/>

References

- [1] GAJENTAAN, Anka ; OVERMARS, Mark H.: On a class of $O(n^2)$ problems in computational geometry. In: *Computational geometry* 5 (1995), Nr. 3, S. 165–185
- [2] BAREQUET, Gill ; HAR-PELED, Sarel: Polygon Containment and Tranlational In-Hausdorf-Distance between Segment Sets are 3Sum-Hard. In: *International Journal of Computational Geometry & Applications* 11 (2001), Nr. 04, S. 465–474
- [3] DEMAINE, Erik D. ; MITCHELL, Joseph S. B. ; O’ROURKE, Joseph: *The Open Problems Project*. <http://cs.smith.edu/~jorourke/TOPP/>, 09 2017. – (Accessed on 2020/07/30)
- [4] DEMAINE, Erik D. ; MITCHELL, Joseph S. B. ; O’ROURKE, Joseph: *The Open Problems Project - Problem 11: 3SUM Hard Problems*. <http://cs.smith.edu/~jorourke/TOPP/P11.html#Problem.11>, 09 2017. – (Accessed on 2020/07/30)
- [5] DEMAINE, Erik D. ; MITCHELL, Joseph S. B. ; O’ROURKE, Joseph: *The Open Problems Project - Problem 41: Sorting $X + Y$ (Pairwise Sums)*. <http://cs.smith.edu/~jorourke/TOPP/P41.html#Problem.41>, 09 2017. – (Accessed on 2020/07/30)
- [6] ERICKSON, Jeff u. a.: Lower Bounds for Linear Satisfiability Problems. In: *SODA*, 1995, S. 388–395
- [7] EDELSBRUNNER, Herbert ; O’ROURKE, Joseph ; SEIDEL, Raimund: Constructing arrangements of lines and hyperplanes with applications. In: *SIAM Journal on Computing* 15 (1986), Nr. 2, S. 341–363
- [8] BARAN, Ilya ; DEMAINE, Erik D. ; PĂTRAȘCU, Mihai: Subquadratic algorithms for 3SUM. In: *Workshop on Algorithms and Data Structures* Springer, 2005, S. 409–421
- [9] AILON, Nir ; CHAZELLE, Bernard: Lower bounds for linear degeneracy testing. In: *Journal of the ACM (JACM)* 52 (2005), Nr. 2, S. 157–171
- [10] GRØNLUND, Allan ; PETTIE, Seth: *Threesomes, Degenerates, and Love Triangles*. 2014
- [11] FREUND, Ari: Improved subquadratic 3SUM. In: *Algorithmica* 77 (2017), Nr. 2, S. 440–458
- [12] GOLD, Omer ; SHARIR, Micha: Improved Bounds for 3SUM, k-SUM, and Linear Degeneracy. In: PRUHS, Kirk (Hrsg.) ; SOHLER, Christian (Hrsg.): *25th Annual European Symposium on Algorithms (ESA 2017)* Bd. 87. Dagstuhl, Germany : Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2017 (Leibniz International Proceedings in Informatics (LIPIcs)). – ISBN 978–3–95977–049–1, 42:1–42:13
- [13] CHAN, Timothy M.: More logarithmic-factor speedups for 3SUM,(median,+)-convolution, and some geometric 3SUM-hard problems. In: *ACM Transactions on Algorithms (TALG)* 16 (2019), Nr. 1, S. 1–23
- [14] KANE, Daniel M. ; LOVETT, Shachar ; MORAN, Shay: *Near-optimal linear decision trees for k-SUM and related problems*. 2017
- [15] KOPELOWITZ, Tsvi ; PETTIE, Seth ; PORAT, Ely: *Higher Lower Bounds from the 3SUM Conjecture*. 2014