A short article on the Subset Sum Problem

Jan Olucha Fuentes

June 2022

Abstract

In this paper we prove a sufficient condition to show the Subset Sum Problem has a solution for a given configuration.

1 Introduction

The Subset Sum Problem (SSP) is a problem in Mathematics and Computer Science that asks the following question:
Let $U \subset \mathbb{N}$ be a non-empty subset of the natural numbers. Let $t \in \mathbb{N}$ be a natural number. Is it possible to construct t by only performing addition on the elements of U?
In other words, does there exist a subset $V \subset U$ with $\sum_i V_i = t$? The computational complexity of this problem is known to be exponential for the best algorithms.
In this paper we study a sufficient but not necessary condition to show if there exists a solution.

2 Definitions and Lemmas

Definition 2.1: The *span* of a set of natural numbers $S \subset \mathbb{N}$ is the set of all numbers that can be constructed by performing addition on the elements of S. For example: let $S = \{1, 5, 7\}$, then $span(S) = \{1, 5, 6, 7, 8, 12, 13\}$.

Definition 2.2: A set $S \subset \mathbb{N}$ is *spanning* $\iff \forall n \in [1, \sum_i S_i], n \in span(S)$. For example: $S = \{1, 2, 3, 4\}$ is spanning since $span(S) = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

Definition 2.3: Natural numbers $a, b, c \in \mathbb{N}$ are *independent* if and only if $c \notin span([a, b])$

Now the question arises as to how we may construct spanning sets.

Lemma 2.4 Let $S \subset \mathbb{N}$ be a spanning set, let c be an integer which is *not independent* from S then the set $S|c$ (S adjoin c) is still spanning, this is because we are not adding an element that was outside the span of S in the first place, so if c was spanned before, it is still spanned.

Lemma 2.5: The set $S = [1, 2, ..., n] \subset \mathbb{N}$, is spanning
Proof: The proof is by induction:

> **Case n=2**: We see that $span(S) = [1, 2, 3]$
> **Inductive hypothesis:** Assume the set $S = [1, 2, ..., n-1]$ is spanning
>
> Now consider the set $S' = [1, 2, ..., n]$, note that $n \in span(S)$, this is because $n \leq \sum_{i=1}^{n-1} i$ $\forall n \geq 3$. Using **Lemma 2.4** we see S' is spanning. ■

We now have proved the rather unsurprising result that the natural numbers from 1 to n span all the possible numbers between 1 and $\sum_{i=1}^{n} i$. But we now ask ourselves: Can we do better?
The answer turns out to be yes. We now introduce the idea of a *minimal spanning set*

Definition 2.6: A *minimal spanning set* of a number $n \in \mathbb{N}$ is the set with the least cardinality that spans all the numbers between 1 and n. For example: the minimal spanning set of 7 is [1,2,4]. Note these sets are not unique, for the minimal spanning sets of 6 are [1,2,3] and [1,2,4], however for reasons that will become obvious now, we will say that the minimal spanning set of 6 is [1,2,4]

Lemma 2.7: The minimum spanning set of a number n is [1,2,4,...,$2^{\lfloor \log_2(n) \rfloor}$]

Proof: The construction of this set becomes clear if we consider the binary representation of n.

> **Recall:** A binary representation of a natural number n is a vector $\vec{B} \in \mathbb{Z}_2^{\lfloor \log_2(n) \rfloor}$ where one has $n = \sum_i \vec{B}_i 2^i$

We see that with $\lfloor \log_2(n) \rfloor$ bits, we may represent every number between 1 and $2^n - 1$ which is precisely the maximum number that can be spanned with [1,2,4,...,$2^{\lfloor \log_2(n) \rfloor}$]. Hence we see it is spanning. Moreover, if we still want to be able to represent every number in that range, we cannot remove any power of two from the set, hence it is minimal.

Now we look back to the original Subset Sum Problem and try to apply what we have learnt:

Corollary 2.8: Let $U \subset \mathbb{N}$ be a non-empty subset of the natural numbers. Let $t \in \mathbb{N}$ be a natural number.

[1,2,4,...,$2^{\lfloor \log_2(t) \rfloor}$]$\subset U \implies \exists\, a_1, a_2, ..., a_n \in U$ with $a_1 + a_2 + ... + a_n = t$

In other words, if the minimal spanning set of t is in U, then t can be constructed.

Proof: This directly follows from **Lemma 2.7**, combined with the fact that if a set U is a subset of a set S, then the span of U is a subset of the span of S.■

Thus, by **Corollary 2.8** it is a sufficient condition to check that [1,2,4,...,$2^{\lfloor \log_2(t) \rfloor}$]$\subset U$

3 Final remarks:

In this paper we have proved a sufficient condition to show if the Subset Sum Problem can be solved for a given set U of size m and a given target t. This condition is not necessary however. This condition may be beneficial for helping to solve the problem when a large set is given, because regardless of the size of the input set, the computational complexity of this approach is $O(\log_2(t))$ so checking the aforementioned condition should not take an excessive amount of computations and if satisfied, finishes the problem immediately.