

Problém SUBSET-SUM

Vstup: Sekvence přirozených čísel a_1, a_2, \dots, a_n a přirozené číslo s ?

Otázka: Existuje množina $I \subseteq \{1, 2, \dots, n\}$ taková, že $\sum_{i \in I} a_i = s$?

Jinak řečeno, ptáme se zda z dané (multi)množiny čísel je možné vybrat podmnožinu, jejíž součet je s .

Příklad: Pro vstup tvořený čísly 3, 5, 2, 3, 7 a číslem $s = 15$ je odpověď **ANO**, neboť $3 + 5 + 7 = 15$.

Pro vstup tvořený čísly 3, 5, 2, 3, 7 a číslem $s = 16$ je odpověď **NE**, neboť žádná podmnožina těchto čísel nedává součet 16.

Poznámka:

Pořadí čísel a_1, a_2, \dots, a_n na vstupu není důležité.

Všimněte si však určitého rozdílu oproti tomu, kdybychom problém formulovali tak, že vstupem je množina $\{a_1, a_2, \dots, a_n\}$ a číslo s :

V množině se čísla neopakují, zatímco v sekvenci se může totéž číslo vyskytnout vícekrát.

Problém SUBSET-SUM je speciálním případem **problému batohu** (knapsack problem):

Knapsack problem

Vstup: Sekvence dvojic přirozených čísel $(a_1, b_1), (a_2, b_2), \dots, (a_n, b_n)$ a dvě přirozená čísla s a t .

Otázka: Existuje množina $I \subseteq \{1, 2, \dots, n\}$ taková, že $\sum_{i \in I} a_i \leq s$ a $\sum_{i \in I} b_i \geq t$?

Neformálně můžeme problém batohu formulovat takto:

Máme n předmětů, kde i -tý předmět váží a_i gramů a má cenu b_i Kč. Do batohu se vejdou předměty o maximální celkové váze s gramů.

Otázka zní, zda můžeme z předmětů vybrat podmnožinu, která by se vešla do batohu a měla celkovou cenu alespoň t Kč.

Poznámka:

Zde jsme problém batohu formulovali jako rozhodovací problém. Běžnější je formulovat tento problém jako optimalizační problém, kde je cílem najít takovou množinu $I \subseteq \{1, 2, \dots, n\}$, kde hodnota $\sum_{i \in I} b_i$ je maximální, přičemž ovšem musí být dodržena podmínka $\sum_{i \in I} a_i \leq s$, tj. vybrat předměty s maximální celkovou cenou tak, aby nebyla překročena kapacita batohu.

Převod 3-SAT na SUBSET-SUM

To, že SUBSET-SUM je speciálním případem problému batohu, vidíme z následující (téměř triviální) redukce:

Řekněme, že $a_1, a_2, \dots, a_n, s_1$ je instance problému SUBSET-SUM. Je očividné, že pro instanci problému batohu, kde máme sekvenci $(a_1, a_1), (a_2, a_2), \dots, (a_n, a_n)$, $s = s_1$ a $t = s_1$, je odpověď stejná jako pro původní instanci SUBSET-SUM.

Převod 3-SAT na SUBSET-SUM

Pokud chceme studovat složitost problémů jako jsou SUBSET-SUM nebo problém batohu, je dobré si nejprve ujasnit, co považujeme za velikost vstupu.

Asi nejpřirozenější je definovat velikost vstupu jako celkový počet bitů, který potřebujeme k zápisu instance.

Musíme však určit, jakým způsobem jsou na vstupu zadána přirozená čísla – zda binárně (případně v jiné číselné soustavě o základu alespoň 2, např. desítkové nebo šestnáctkové) nebo unárně.

V dalším výkladu budeme předpokládat, že čísla jsou na vstupu zadána **binárně**, tj. že velikost vstupu je úměrná součtu délek binárních zápisů jednotlivých čísel na vstupu.

Není těžké se přesvědčit, že SUBSET-SUM i problém batohu (jeho rozhodovací varianta) patří do třídy **NPTIME**:

- Nedeterministický algoritmus nejprve nedeterministicky zvolí podmnožinu prvků sekvence na vstupu a poté (deterministicky) ověří, zda splňuje splňuje danou podmínku (resp. podmínky).

Je zřejmé, že toto ověření je možné provést v čase polynomiálním vzhledem k velikosti instance.

Ukážeme si, že problém SUBSET-SUM (a tím pádem i problém batohu) je **NP-těžký**.

Převod 3-SAT na SUBSET-SUM

NP-obtížnost problému SUBSET-SUM ukážeme pomocí redukce ze známého **NP**-úplného problému 3-SAT:

3-SAT

Vstup: Booleovské formule φ v konjunktivní normální formě, kde každá klauzule obsahuje právě tři literály.

Otázka: Je formule φ splnitelná?

Převod 3-SAT na SUBSET-SUM

Předpokládejme, že máme danu instanci problému 3-SAT, tj. formuli φ tvaru

$$C_1 \wedge C_2 \wedge \cdots \wedge C_k$$

kde C_1, C_2, \dots, C_k jsou jednotlivé **klauzule**, přičemž každá klauzule C_j je tvaru

$$(L_{j,1} \vee L_{j,2} \vee L_{j,3})$$

kde $L_{j,1}, L_{j,2}, L_{j,3}$ jsou jednotlivé **literály** v této klauzuli, přičemž každý z těchto literálů je buď tvaru x_i nebo tvaru $\neg x_i$, kde x_i je nějaká booleovská **proměnná**.

Předpokládejme dále, že množina všech proměnných, které se vyskytují ve formuli φ , je

$$\{x_1, x_2, \dots, x_m\}$$

Příklad: Vezměme si třeba následující formuli φ :

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Tato formule obsahuje klauzule

$$C_1 = (x_1 \vee \neg x_2 \vee x_3)$$

$$C_2 = (x_2 \vee \neg x_3 \vee x_4)$$

$$C_3 = (x_1 \vee \neg x_3 \vee \neg x_4)$$

$$C_4 = (\neg x_1 \vee \neg x_2 \vee x_4)$$

Klauzule $(x_1 \vee \neg x_3 \vee \neg x_4)$ obsahuje literály x_1 , $\neg x_3$ a $\neg x_4$.

Množina všech proměnných, které se vyskytují ve formuli φ je $\{x_1, x_2, x_3, x_4\}$.

Převod 3-SAT na SUBSET-SUM

K dané formuli φ sestrojíme instanci problému SUBSET-SUM, která bude obsahovat multimnožinu čísel X a číslo s .

Čísla v multimnožině X i číslo s budeme popisovat tak, že popíšeme, jak budou vypadat jednotlivé číslice daného čísla zapsaného v číselné soustavě o základu $d = 10$.

Poznámka: Číselný základ $d = 10$ jsme zvolili pro větší názornost, konstrukce bude stejně dobře fungovat i s libovolným jiným základem $d \geq 7$.

Převod 3-SAT na SUBSET-SUM

Ještě si připomeňme, že $b_k b_{k-1} b_{k-2} \cdots b_1 b_0$ je v číselné soustavě o základu d zápisem čísla

$$\sum_{i=0}^k b_i \cdot d^i$$

V tomto zápise b_k, b_{k-1}, \dots, b_0 reprezentují jednotlivé číslice, přičemž pro všechna b_i platí $0 \leq b_i < d$.

Budeme-li dále mluvit o číslici na pozici i , máme tím na mysli hodnotu b_i ve výše uvedeném zápise.

Příklad: V čísle 732594 je na pozici 5 číslice 7, na pozici 4 číslice 3 a na pozici 0 číslice 4.

Převod 3-SAT na SUBSET-SUM

Připomeňme, že předpokládáme, že formule φ obsahuje proměnné x_1, x_2, \dots, x_m a skládá se z klauzulí C_1, C_2, \dots, C_k .

V instanci problému SUBSET-SUM, kterou budeme vytvářet, budeme jednotlivá čísla zapisovat jako $(m + k)$ -místná čísla v číselné soustavě o základu $d = 10$ s tím, že:

- Každé proměnné x_i přidělíme jednu pozici v těchto číslech (každé jinou).
- Každé klauzuli C_j přidělíme jednu ze zbylých pozic (každé jinou).

Jak konkrétně toto přiřazení provedeme, není pro další konstrukci příliš podstatné.

Jednou z možností je například:

- Přiřadit proměnným x_1, x_2, \dots, x_m pozice $0, 1, \dots, m - 1$ (tj. proměnné x_i je přiřazena pozice $i - 1$).
- Přiřadit klauzulím C_1, C_2, \dots, C_k pozice $m, m + 1, \dots, m + k - 1$ (tj. klauzuli C_j je přiřazena pozice $j + m - 1$).

V dalším výkladu budeme tedy hovořit o pozici odpovídající proměnné x_i či o pozici odpovídající klauzuli C_j .

Převod 3-SAT na SUBSET-SUM

Pro přehlednější výklad zavedeme následující značení:

Řekněme, že y je nějaké číslo, a že toto číslo zapíšeme v číselné soustavě o základu d .

- $y[x_i]$ bude označovat číslici na pozici odpovídající proměnné x_i .
- $y[C_j]$ bude označení číslici na pozici odpovídající klauzuli C_j .

Ve výše popsaném konkrétním přiřazení pozic bude tedy platit, že číslo y bude zapsáno jako

$$y[C_k] y[C_{k-1}] \cdots y[C_1] y[x_m] y[x_{m-1}] \cdots y[x_1]$$

neboli

$$y = \sum_{i=1}^m y[x_i] \cdot d^{i-1} + \sum_{j=1}^k y[C_j] \cdot d^{j+m-1}$$

Převod 3-SAT na SUBSET-SUM

Nyní popíšeme, jak vytvořit instanci problému SUBSET-SUM s multimnožinou X a číslem s :

- Pro každou proměnnou x_i z množiny $\{x_1, x_2, \dots, x_m\}$ přidáme do X dvojici čísel y_i a y'_i , jejichž číslce jsou definovány následovně:
 - $y_i[x_i] = y'_i[x_i] = 1$
 - Pro $x_j \neq x_i$ je $y_i[x_j] = y'_i[x_j] = 0$
 - Pro každou klauzuli C_j je hodnota $y_i[C_j]$ rovna počtu výskytů literálu x_i v klauzuli C_j .
 - Pro každou klauzuli C_j je hodnota $y'_i[C_j]$ rovna počtu výskytů literálu $\neg x_i$ v klauzuli C_j .

Převod 3-SAT na SUBSET-SUM

- Pro každou klauzuli C_j z množiny $\{C_1, C_2, \dots, C_k\}$ přidáme do X dvojici čísel z_j a z'_j , jejichž číslice jsou definovány následovně:
 - $z_j[C_j] = 1$
 - $z'_j[C_j] = 2$
 - Všechny ostatní číslice v z_j i z'_j jsou 0.
- Vytvoříme hodnotu s , kde:
 - $s[x_i] = 1$ pro všechny proměnné x_i .
 - $s[C_j] = 4$ pro všechny klauzule C_j .

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1		

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C₄	C ₃	C ₂	C ₁	x ₄	x ₃	x ₂	x ₁	
y ₁	0	1	0	1	0	0	0	1	x ₁
y' ₁	1	0	0	0	0	0	0	1	¬x ₁
y ₂	0	0	1	0	0	0	1	0	x ₂
y' ₂	1	0	0	1	0	0	1	0	¬x ₂
y ₃	0	0	0	1	0	1	0	0	x ₃
y' ₃	0	1	1	0	0	1	0	0	¬x ₃
y ₄	1	0	1	0	1	0	0	0	x ₄
y' ₄	0	1	0	0	1	0	0	0	¬x ₄
z ₁	0	0	0	1	0	0	0	0	C ₁
z' ₁	0	0	0	2	0	0	0	0	C ₁
z ₂	0	0	1	0	0	0	0	0	C ₂
z' ₂	0	0	2	0	0	0	0	0	C ₂
z ₃	0	1	0	0	0	0	0	0	C ₃
z' ₃	0	2	0	0	0	0	0	0	C ₃
z ₄	1	0	0	0	0	0	0	0	C ₄
z' ₄	2	0	0	0	0	0	0	0	C ₄
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Převod 3-SAT na SUBSET-SUM

Příklad instance problému SUBSET-SUM odpovídající níže uvedené formuli:

	C_4	C_3	C_2	C_1	x_4	x_3	x_2	x_1	
y_1	0	1	0	1	0	0	0	1	x_1
y_1'	1	0	0	0	0	0	0	1	$\neg x_1$
y_2	0	0	1	0	0	0	1	0	x_2
y_2'	1	0	0	1	0	0	1	0	$\neg x_2$
y_3	0	0	0	1	0	1	0	0	x_3
y_3'	0	1	1	0	0	1	0	0	$\neg x_3$
y_4	1	0	1	0	1	0	0	0	x_4
y_4'	0	1	0	0	1	0	0	0	$\neg x_4$
z_1	0	0	0	1	0	0	0	0	C_1
z_1'	0	0	0	2	0	0	0	0	C_1
z_2	0	0	1	0	0	0	0	0	C_2
z_2'	0	0	2	0	0	0	0	0	C_2
z_3	0	1	0	0	0	0	0	0	C_3
z_3'	0	2	0	0	0	0	0	0	C_3
z_4	1	0	0	0	0	0	0	0	C_4
z_4'	2	0	0	0	0	0	0	0	C_4
s	4	4	4	4	1	1	1	1	

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

Výsledná instance vytvořená k formuli

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee \neg x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee \neg x_2 \vee x_4)$$

bude tedy vypadat takto:

$$X = \{1010001, 1000001, 100010, 10010010, 10100, 1100100, \\ 10101000, 1001000, 10000, 20000, 100000, 200000, 1000000, \\ 2000000, 10000000, 20000000\}$$

$$s = 44441111$$

Převod 3-SAT na SUBSET-SUM

Všimněme si ještě, že výše popsanou konstrukci můžeme stručně popsat pomocí následujících vztahů:

$$y_i = d^{i-1} + \sum_{j=1}^k c(i, j) \cdot d^{j+m-1} \quad y'_i = d^{i-1} + \sum_{j=1}^k c'(i, j) \cdot d^{j+m-1}$$

$$z_j = d^{j+m-1} \quad z'_j = 2 \cdot d^{j+m-1}$$

$$s = \sum_{i=1}^m d^{i-1} + \sum_{j=1}^k 4 \cdot d^{j+m-1}$$

kde $c(i, j)$ udává počet výskytů literálu x_i v klauzuli C_j a $c'(i, j)$ počet výskytů literálu $\neg x_i$ v klauzuli C_j .

Je zřejmé, že konstrukci je možné provést v polynomiálním čase.

Zbývá ukázat, že konstrukce je korektní, tedy, že z X je možné vybrat podmnožinu, jejíž prvky dávají součet s , právě tehdy, když je původní formule φ splnitelná.

Převod 3-SAT na SUBSET-SUM

Než přistoupíme k vlastnímu důkazu korektnosti, všimněme si nejprve, jak vypadají jednotlivé sloupce číslic, když zapíšeme čísla v X pod sebou:

- Ve sloupcích odpovídajících proměnným x_i jsou vždy právě dvě jedničky (v y_i a y_i').
- Ve sloupcích odpovídajících klauzulím C_j je součet v daném sloupci přes všechna y_i a y_i' vždy roven 3 (tj. je to celkový počet literálů v klauzuli C_j), a dále z_j a z_j' obsahují vždy číslice 1 a 2
- Všechny ostatní hodnoty jsou 0.

Vidíme, že součet v každém sloupci je maximálně 6, a že vzhledem k tomu, že $d \geq 7$, nikdy při součtu žádné podmnožiny nedojde k přenosu z jednoho sloupce do druhého.

Předpokládejme nejprve, že formule φ je splnitelná.

Pak tedy existuje nějaké ohodnocení booleovských proměnných ν , při kterém $[\varphi]_\nu = \text{TRUE}$.

Budeme vytvářet (multi)množinu $Y \subseteq X$ takovou, aby součet hodnot v Y byl roven s :

- Pro každou proměnnou x_i takovou, že $[x_i]_\nu = \text{TRUE}$, přidáme do Y číslo y_i .
- Pro každou proměnnou x_i takovou, že $[x_i]_\nu = \text{FALSE}$, přidáme do Y číslo y'_i .

Převod 3-SAT na SUBSET-SUM

Když není sečteme hodnoty čísel, které jsme zatím přidali do Y , budou součty v jednotlivých sloupcích vypadat takto:

- Ve sloupcích odpovídajících proměnným x_i bude součty vždy 1 , neboť do Y jsme dali vždy právě jednu z hodnot y_i, y_i' .
V těchto sloupcích tedy součet souhlasí s požadovanou hodnotou s .
- Ve sloupcích odpovídajících klauzulím C_j bude vždy součet roven počtu literálů v dané klauzuli, které mají při ohodnocení ν hodnotu **TRUE**.

Při ohodnocení ν pro všechny klauzule C_j platí, že $[C_j]_\nu = \mathbf{TRUE}$ a tedy alespoň jeden literál v klauzuli C_j musí mít při ohodnocení ν hodnotu **TRUE**.

Součet v daném sloupci tedy bude číslo v intervalu 1 až 3 .

Převod 3-SAT na SUBSET-SUM

Pro každý sloupec odpovídající nějaké klauzuli C_j přidáme do Y jedno nebo případně obě z čísel z_j, z'_j tak, abychom dorovnali součet v tomto sloupci na 4.

(Všiměte si, že vzhledem k tomu, že součet je v intervalu 1 až 3, tak je toto dorovnání na součet 4 vždy možné.)

Čísla z_j i z'_j mají na všech ostatních pozicích hodnoty 0, takže součty v ostatních sloupcích neovlivní.

Vidíme tedy, že pokud je formule φ splnitelná, je možné z X vybrat podmnožinu, která dává součet s .

Předpokládejme nyní, že existuje nějaká (multi)množina $Y \subseteq X$, jež dává součet s .

Je zřejmé, že pro každé $i \in \{1, 2, \dots, m\}$ musí Y obsahovat právě jednu z hodnot y_i, y_i' .

Zvolme tedy následující ohodnocení ν :

- Pokud Y obsahuje y_i , položme $[x_i]_\nu = \text{TRUE}$.
- Pokud Y obsahuje y_i' , položme $[x_i]_\nu = \text{FALSE}$.

Podobně jako v předchozím případě udávají součty ve sloupcích odpovídajících klauzulím po odečtení případných z_j a z_j' , která se mohou v Y vyskytovat, počty literálů v dané klauzuli C_j , které mají při ohodnocení ν hodnotu **TRUE**.

Převod 3-SAT na SUBSET-SUM

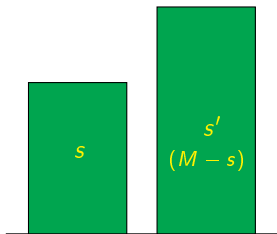
Pro každou klauzuli C_j musí být tento počet roven mimálně 1, protože pokud by byl 0, nebylo by ho možné pomocí z_i a z'_i dorovnat na 4.

V libovolné klauzuli C_j má tedy při ohodnocení ν alespoň jeden literál hodnotu **TRUE**, takže $[\varphi]_\nu = \mathbf{TRUE}$ a φ je tedy splnitelná.

Tím je důkaz korektnosti hotov.

Převod 3-SAT na SUBSET-SUM

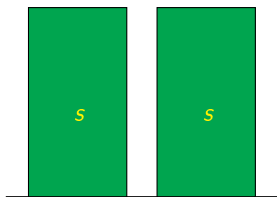
V problému SUBSET-SUM jde o to, rozdělit čísla v dané multimnožině do dvou podmnožin takových, že součet čísel v jedné je s a součet v druhé je $s' = M - s$, kde M je součet všech čísel v dané multimnožině.



Je očividné, že podmnožinu se součtem s je možné vybrat právě tehdy, když je možné vybrat podmnožinu se součtem s' .

Převod 3-SAT na SUBSET-SUM

Speciálním případem problému SUBSET-SUM je případ, kdy $s = s'$, kdy je cílem rozdělit zadanou multimnožinu do dvou podmnožin se stejně velkým součtem:



SUBSET-SUM-1/2

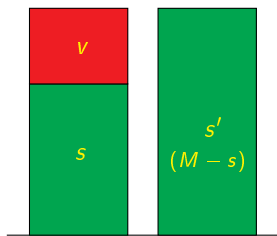
Vstup: Sekvence přirozených čísel a_1, a_2, \dots, a_n .

Otázka: Existuje množina $I \subseteq \{1, 2, \dots, n\}$ taková, že

$$\sum_{i \in I} a_i = \frac{1}{2} \sum_{i=1}^n a_i ?$$

Převod 3-SAT na SUBSET-SUM

I tento problém SUBSET-SUM-1/2 je NP-těžký, jak je vidět z následující snadné redukce ze SUBSET-SUM:



Do multimnožiny X přidáme nový prvek $v = s' - s$ (pokud $s < s'$) nebo $v = s - s'$ (pokud $s > s'$).
(Pokud $s = s'$, nemusíme dělat nic.)

Převod 3-SAT na SUBSET-SUM

Je zřejmé, že pokud multimnožinu X rozšířenou o nový prvek v rozdělíme na dvě podmnožiny se stejným součtem, musí tento prvek padnout do jedné z těchto podmnožin, a po jeho odebrání nám zbydou podmnožiny se součty s a s' .

Vidíme tedy, že řešení takto vytvořené instance SUBSET-SUM-1/2 existuje právě tehdy, když existuje řešení původní instance SUBSET-SUM.