about this note

This is a note by team 'seed71' on Kaggle competition 'Santa Gift Matching Challenge'. In this note we prove that the maxmum value possible is 0.9363015472581603694371374 74.

First of all, let us introduce some notations.
Let $\mathcal{M}$ denote all possible matchings such that every Triplets and Twins are given the same gift, and let $\mathcal{M}'$ denote all matchings.
Let $CH(m)$ denote sum of $6 \times ChildHappiness$ and let $SH(m)$ denote sum of $6 \times GiftHappiness$, where $m \in \mathcal{M}'$ is a matching. We multiply by 6 so that these values will be integers.

The goal is to find $m \in \mathcal{M}$ that maximize $S(m) := \{10 \times CH(m)\}^3 + \{SH(m)\}^3$. (Note that $10\times$ comes from the fact that $MaxGiftHappiness = 2000$ while $MaxChildHappiness = 200$). Our goal is to prove the following since we have such a matching.

Theorem 1.

$\max_{m \in \mathcal{M}} S(m)$ is attained when $CH(m) = 1173959622$ and $SH(m) = 1703388$.

To prove this, we first show $CH(m)$ is maximized.

Proposition 2.

$\max_{m \in \mathcal{M}} CH(m) \leq 1173959622$

Lemma 3.

$\max_{m \in \mathcal{M}'} CH(m) = 1173959626$

proof.
We can obtain this by solving a min-cost max-flow problem (hoge1.py). □

proof of Proposition 2.
Since $\mathcal{M} \subset \mathcal{M}'$, $\max_{m \in \mathcal{M}} CH(m) \leq 1173959626$ is obvious.
On the other hand, $CH(m)$ must be a multiple of 6 when $m \in \mathcal{M}$ while $1173959626 \equiv 4 (\bmod 6)$, thus we obtain the desired inequality. □

As a result of many trial and error, we found it very hard to please the twins [34267, 34268]. The following proposition states we have to assign gift 207 to them if we want to maximize $CH(m)$.

Proposition 4.

Let $\mathcal{M}'_j \subset \mathcal{M}'$ ($j \in \{0, 1, 2, \ldots, 999\}$) denote the set of matchings where child 34267 and child 34268 are assigned gift j, then

$\max_{m \in \mathcal{M}'_j} CH(m) = 1173959622$ if and only if $j = 207$.

proof.
Obtained by the brute force search (hoge2.py), where the following lemma helps us reduce the search range. □

Lemma 5.

If $ChildHappiness$ for child 34267 is -1 when assigned gift j, then $\max_{m \in \mathcal{M}'_j} CH(m) \le 1173959620$.

proof.
By solving a problem to maximize the sum of $6 \times ChildHappiness$ where the twins [34267, 34268] are ignored (hoge3.py), we see the maximum is 1173959632. □

Next, we try to maximize $SH(m)$ in condition that $CH(m) = 1173959622$.

Proposition 6.

$\max_{m \in \cup_j \mathcal{M}'_j, CH(m)=1173959622} SH(m) = 1703388$.

proof.
We only need to search the case when $j = 207$ (Propsition 4.). Let us consider a problem to maximize $10000 \times CH(m) + SH(m)$ where $m \in \mathcal{M}'_{207}$. Again, we can solve this problem as a min-cost max-flow problem (hoge4.py), the maximum is attained when $CH(m) = 1173959622, SH(m) = 1703388$. This means $m \in \cup_j \mathcal{M}'_j$ with $CH(m) = 1173959622$ and $SH(m) > 1703388$ does not exist. □

Corollary 7.

$\max_{m \in \mathcal{M}, CH(m)=1173959622} SH(m) \le 1703388$.

proof.
Obvious because $\mathcal{M} \subset \cup_j \mathcal{M}'_j$ □

Together with the matching we constarcted, we have shown that $\max_{m \in \mathcal{M}} CH(m) = 1173959622$ and that $\max_{m \in \mathcal{M}, CH(m)=1173959622} SH(m) = 1703388$. We set $S_M := (10 \times 1173959622)^3 + 1703388^3$ for convenience. This looks like the maximum. What is left to show is that $m \in \mathcal{M}$ such that $CH(m)$ is smaller cannot have much larger $SH(m)$ so that $S(m)$ exceed S_M.

Proposition 8.

$\max_{m \in \mathcal{M}', CH(m)<1173959622} S(m) \le S_M$.

proof.

We prove this by covering the area $\{CH(m), SH(m)\}_{m \in \mathcal{M}'}$.

First, let us solve a problem to maximize $10000 \times CH(m) + 2 \times SH(m)$ where $m \in \mathcal{M}'$. This problem is also a min-cost max-flow problem (hoge5.py). The maximum is attained when $CH(m) = 1173959622, SH(m) = 1709307$. This shows, if $CH(m) = 1173959622 - n$, then $SH(m) \leq 1709307 + 5000 \times n$. By simple calculation we get $\{10 \times (1173959622 - n)\}^3 + (1709307 + 5000 \times n)^3 \leq S_M$ when $1 \leq n \leq 181341$, and hence we have checked $\max_{m \in \mathcal{M}', 1173778280 < CH(m) < 1173959622} S(m) \leq S_M$ (A).

Next, let us solve a problem to maximize $100 \times CH(m) + SH(m)$ where $m \in \mathcal{M}'$. The maximum is attained when $CH(m) = 1173783839, SH(m) = 33226746$, and, in the same way as above, we see $\max_{m \in \mathcal{M}', 1111594258 < CH(m) < 11737838390} S(m) \leq S_M$ (B).

Finaly, let us see the case when $CH(m)$ is small. From the problem settings, it is clear that $\max_{m \in \mathcal{M}'} SH(m) \leq 6006000000$ (the case that 1000 GiftGoodKidsLists have no duplication). Since 6006000000^3 is small compared to S_M, we have a constant $((S_M - 6006000000^3) / 10^3)^{1/3} = 1119029885.25\ldots$ and we have $\max_{m \in \mathcal{M}', CH(m) \leq 1119029885} S(m) \leq S_M$ (C).

Combining (A), (B), and (C), we have the desired inequality. □

Corollary 9.

$\max_{m \in \mathcal{M}, CH(m) < 1173959622} S(m) \leq S_M$.

proof.

Obvious because $\mathcal{M} \subset \mathcal{M}'$ □

Now, it is the time to finish this note, thank you for reading this.

*proof of Theorem*1.

Obvious from Proposition 2. Corollary 7. and Corollary 9. □