Solution to Technical Problem I

Chengyuan Ma

January 17, 2020

0 Result

0.1 Short answer

$\sup\{\Delta\} = 100°\mathrm{C}$.

0.2 Longer answer

$\forall \varepsilon \in (0, 200)$, there exists a procedure consisting of finite number of operations that eventually yields $\Delta = (100 - \varepsilon)°\mathrm{C}$, and there exists no procedure that can possibly produce a higher Δ. The optimal procedure is described in section 3.2 and its proof in section 4.

Contents

1 Assumptions

This solution is based on the following assumptions:

1. The blocks are infinitely divisible. That is, the microscopic structure of matters are completely ignored.

2. No heat loss during heat exchanges, as is stated in the problem.

3. Heat transfers are instant and perfectly uniform, as is stated in the problem as well.

4. Heat transfers can be perfectly modeled by the specific heat equation $\Delta Q = cm\Delta T$.

5. Though not explicitly stated in the problem, it can be inferred from the example given that both blocks have **the same specific heat capacity**. Therefore, **the specific heat capacity term c is omitted in all equations throughout the solution**, because they will be canceled out in the result anyway (alternative interpretation: the system of measurement is chosen in such a way that the value of c is always 1).

6. Degrees Celsius is used as the unit of temperature, and kilogram is used as the unit of mass. For simplicity, **the symbols of these units will not appear in equations below**.

2 Macroscopic Analysis

Before going deeper into specific operations, I analyzed the problem in a macroscopic point of view. That is, I defined some physical quantities of a *system* that can reflect the status of its *elements* to a certain extent. The analysis of these general quantities provided me insights to the design of the optimal procedure.

Note: The term "system" (and "subsystem") is used instead of "object" (and "part") in this section so that these definition and analysis can cover as many cases as possible.

2.1 Defining quantities

Definition 2.1 (Relative internal thermal energy)**.** *The relative internal thermal energy (internal energy for short) of a system is defined as the heat loss needed to cool the whole system down to a specific temperature (*$0°\mathrm{C}$ *here, this is why it is called "relative"):*

$$U = \int T \mathrm{d}m \tag{1}$$

In the integration, T denotes the temperature of the differential mass.

Definition 2.2 (Average temperature)**.** *It follows directly that the average temperature of a system with total mass m is:*

$$\overline{T} = \frac{U}{m} \tag{2}$$

Remark. *If we see the red and the blue block as two systems. Then it follows that*

$$\Delta = B - R = \overline{T_{blue}} - \overline{T_{red}} \tag{3}$$

Since both systems have a total mass of 1*, it further follows that*

$$\Delta = U_{blue} - U_{red} \tag{4}$$

All these physical quantities defined above help reveal how hot, **on average**, a system is. However, **the variance** in temperature distribution is loss when computing these quantities. To account for variance:

Definition 2.3 (Temperature variance)**.** *The temperature variance is defined similar to how variance of a group of data is defined in statistics:*

$$S = \int \left(T - \overline{T}\right)^2 \mathrm{d}m \tag{5}$$

Corollary 2.0.1. $S = 0$ *if and only if all objects in the system have the same uniform temperature.*

Corollary 2.0.2. *The temperature variance of a union of two systems is the sum of the temperature variance of the two systems.*

2.2 Some important properties

Now that these quantities have been defined, it is time to analyze their properties.

Lemma 2.1. *Temperature variance always decreases with heat exchanges within the system.*

Proof. Assume in an heat exchange operation, two mass m_1, m_2 with *uniform* temperature T_1, T_2 exchange heat with each other. The final temperature of the masses is, following either the specific heat equation or the definition of average temperature above:

$$T = \frac{m_1 T_1 + m_2 T_2}{m_1 + m_2} \tag{6}$$

The change of the system's temperature variance can be then expressed by:

$$\begin{aligned}\Delta S &= (m_1 + m_2)\left(\frac{m_1 T_1 + m_2 T_2}{m_1 + m_2} - \overline{T}\right)^2 - \left[m_1(T_1 - \overline{T})^2 + m_2(T_2 - \overline{T})^2\right] \\ &= -\frac{m_1 m_2 (T_1 - T_2)^2}{m_1 + m_2} < 0\end{aligned} \tag{7}$$

Therefore, the temperature variance of the system always decreases with heat exchange operations, which means that the temperature distribution will become more and more uniform overtime (This, in some sense, proves the third law of Thermodynamics in our context).

Note: This, however, is different from stating that S will finally converge to 0! □

Lemma 2.2 (Bound on difference in internal energy)**.** *Assume a closed system of average temperature $\overline{T}$ is split into two **equal-mass** subsystems, each of mass m, with different total internal energy U_1, U_2 respectively. Then*

$$|U_1 - U_2| \leq 2mS \tag{8}$$

where S is the temperature variance of the whole system.

Proof. First, expand the left hand side:

$$\begin{aligned}|U_1 - U_2|^2 &= U_1^2 + U_2^2 - 2U_1U_2 \\ &= \left(\int T_1 \mathrm{d}m_1\right)^2 + \left(\int T_2 \mathrm{d}m_2\right)^2 - 2U_1U_2\end{aligned} \tag{9}$$

(The subscript 1 and 2 in the integrals are used merely to distinguish between the two subsystems)

Then, since $y = x^2$ is convex, apply Jensen's Inequality:

$$\left(\frac{1}{m}\int T \mathrm{d}m\right)^2 \leq \frac{1}{m}\int T^2 \mathrm{d}m \tag{10}$$

The equality holds if and only if T is constant (temperature is uniformly distributed).

$$\begin{aligned}|U_1 - U_2|^2 &\leq m\int T_1^2 \mathrm{d}m_1 + m\int T_2^2 \mathrm{d}m_2 - 2U_1U_2 \\ &= m\int T^2 \mathrm{d}m - 2U_1U_2 \\ &= m\left[\int (T-\overline{T})^2 \mathrm{d}m + \int 2T\overline{T}\mathrm{d}m - \int \overline{T}^2 \mathrm{d}m\right] - 2U_1U_2 \\ &= m\left[S + 2\overline{T}\int T\mathrm{d}m - 2m\overline{T}^2\right] - \frac{1}{2}\left[(U_1+U_2)^2 - (U_1-U_2)^2\right] \\ &= m\left[S + 2\overline{T}(U_1+U_2) - 2m\overline{T}^2\right] - \frac{1}{2}\left[(U_1+U_2)^2 - (U_1-U_2)^2\right] \\ &= m\left[S + 2\overline{T}\left(2m\overline{T}\right) - 2m\overline{T}^2\right] - \frac{1}{2}\left[\left(2m\overline{T}\right)^2 - (U_1-U_2)^2\right] \\ &= mS + \frac{1}{2}(U_1-U_2)^2\end{aligned} \tag{11}$$

Therefore,

$$|U_1 - U_2|^2 \leq 2mS \tag{12}$$

$\square$

Remark. *In the context of this specific problem, the inequality simplifies to:*

$$|\Delta|^2 \leq 2S \tag{13}$$

2.3 Implications

What do these results imply?

First, lemma 2.1 and 2.2 together show that $|\Delta|$ has a decreasing upper bound, and since calculation shows that $2S = \Delta^2 = 10000$ at the beginning, as long as heat exchanges happen $|\Delta| < 100$. This sets **an upper bound on Δ**.

Moreover, in order to maximize Δ (and thus $|\Delta|$), we need to keep its upper bound $2S$ high. This implies that the optimal procedure to maximize Δ should either:

1. have just a few operations,
2. or consist of operations that have little $|\Delta S|$.

This is indeed an insight to how the optimal procedure should be designed, which will be revealed in the next section.

3 Designing the procedure

In this section, a speculative and heuristic procedure will be constructed with insights from the previous section, computer simulation will be implemented and run, and from the output of the simulation the procedure will be speculatively reduced to a form more friendly to mathematical analysis.

Note: The analysis and proof of its optimality are in the next section. This section is more about intuition than logic :)

3.1 Even-division simplification

Though the game rule allows the two blocks to be arbitrarily cut at any time, I expect **all divisions to happen before all heat exchange operations** so that my procedure can focus only on the latter.

Furthermore, I demand that both blocks be initially divided **evenly into n chunks** so that all chunks have the same mass. Because for equal-mass objects, the final temperature after their contact is just the arithmetic mean of their original temperatures, this simplification in fact make further analysis easier.

Moreover, this simplification actually doesn't impair our ability: when n approaches infinity, we can use these tiny chunks to construct objects of various different masses, just like how tiny particles compose various objects in real world. This is exactly why I previously assumed that matters are infinitely divisible and have no microscopic structures: If I want to designate my own "indivisible" matters, I don't expect my "atoms" to be in conflict with any pre-existing physical rules.

3.2 Heuristic procedure

At the end of the last section I concluded that a "good" procedure should either have just a few steps or should have steps of really low $|\Delta S|$. Because I haven't thought of any effective and general procedure that meets the former requirement, minimizing $|\Delta S|$ should be my goal.

Because I adopted the even-division simplification, the expression of ΔS in equation 7 can be simplified to

$$\Delta S \propto (T_1 - T_2)^2 \tag{14}$$

That is, the smaller the temperature difference is, the smaller $|\Delta S|$ will be. This leads to this intuitive procedure:

Procedure 1 Heuristic procedure to maximize Δ by minimizing$|\Delta S|$

Parameter The number of chunks per block $n \in \mathbb{N}^*$

procedure MAXIMIZEΔ
    Divide the red block evenly into n chunks $r_1, \cdots, r_n$
    Divide the blue block evenly into n chunks $b_1, \cdots, b_n$
    loop
        if $\forall i, j = 1, \cdots, n, T(r_i) \leq T(b_j)$ **then** ▷ $T(...)$: Temperature
            break ▷ We can no longer transfer heat from red to blue
        end if
        $i, j \leftarrow \arg\min\{T(r_i) - T(r_j) > 0\}$
        Put chunk r_i and b_j into contact
    end loop
end procedure

This procedure constantly chooses the pair of red and blue chunks with the smallest positive temperature difference and puts them into contact.

It is easy to turn this procedure into a computer simulation program. I ran the program for $n = 2, 3, 4, 5, 50, 2000$ and the here are the results:

Table 1: Result of the heuristic procedure

n	**Final Δ**	**# of operations (contacts)**
2	25	4
3	37.5	9
4	45.3125	16
5	50.7813	25
50	82.0822	2500
2000	97.4770	2607176

Two observations can be made:

1. The resulting Δ approaches 100 as n gets larger. This gives me confidence in the procedure's optimality.

2. The number of heat exchange operations needed is n^2 (the early stop of the program when $n = 2000$ is most possibly due to floating-point round-off errors).

Based on these two operations, and after tracking the intermediate outputs of the program, I was then able to simplify the above the procedure to a more "deterministic" version:

Procedure 2 Simplified heuristic procedure to maximize Δ by minimizing$|\Delta S|$

Parameter The number of chunks per block $n \in \mathbb{N}^*$

procedure MAXIMIZEΔ
  Divide the red block evenly into n chunks $r_1, \cdots, r_n$
  Divide the blue block evenly into n chunks $b_1, \cdots, b_n$
  for $i \in 1, 2, 3, \cdots, n$ **do** ▷ I call this "*a round*"
    for $j \in 1, 2, 3, \cdots, n$ **do**
      Put chunk r_j and b_i into contact
    end for
  end for
end procedure

It becomes obvious in the above procedure that n^2 heat exchange operations are needed. The output of the program is the same as that of procedure 1.

4 Analyzing the procedure

In this section, I will analyze procedure 2 mathematically and prove its optimality. The variables and notations defined in procedure 2 will be used.

Theorem 4.1.

$$\lim_{n \to \infty} \Delta = 100 \tag{15}$$

Proof. Denote the temperature of chunk r_j after round i as $T_{i,j}$, and the operation that puts chunk r_j into contact with chunk b_i in round i as "operation (i,j)".

Boundary conditions: for convenience, let

$$\begin{aligned} &T_{0,j} = 100 \quad \forall j = 1,2,3,\cdots,n \\ &T_{i,0} = 0 \qquad \forall i = 1,2,3,\cdots,n \end{aligned} \tag{16}$$

In each round chunk r_j changes temperature only once. Therefore, the temperature of chunk r_j before operation (i,j) should just be $T_{i-1,j}$, and for chunk b_i, because it has just had contact with chunk r_{j-1}, it should have temperature $T_{i,j-1}$. Thus

$$T_{i,j} = \frac{1}{2}\left(T_{i-1,j} + Ti,j-1\right) \tag{17}$$

Multiplying both sides of the equation with 2^j:

$$2^j T_{i,j} = 2^{j-1} T_{i-1,j} + 2^{j-1} T_{i,j-1} \tag{18}$$

Reducing the recurrence relation to summation and expanding recursively:

$$\begin{aligned} 2^j T_{i,j} &= \sum_{k=1}^{j} 2^{k-1} T_{i-1,k} \\ &= \frac{1}{2}\sum_{k=1}^{j} 2^k T_{i-1,k} \\ &= \frac{1}{2}\sum_{k_i=1}^{j} \frac{1}{2}\sum_{k_{i-1}=1}^{k_i} 2^{k_{i-1}} T_{i-1,k_{i-1}} \\ &= \frac{1}{2^i}\sum_{k_i=1}^{j}\sum_{k_{i-1}=1}^{k_i}\sum_{k_{i-2}=1}^{k_{i-1}}\cdots\sum_{k_1=1}^{k_2} 2^{k_1} T_{0,k_1} \\ &= \frac{1}{2^i}\sum_{j\geq k_i\geq k_{i-1}\geq\cdots\geq k_1\geq 1} 2^{k_1} T_{0,k_1} \end{aligned} \tag{19}$$

Then, we count how much time $2^{k_1} T_{0,k_1}$ appears in the expansion of the summation, which is equal to the number of different assignments to $k_1,\cdots,k_i$ such that $j \geq k_i \geq k_{i-1} \geq \cdots \geq k_1 \geq 1$, which is then equal to the number of ways to partition $j-k_1$ into i non-negative integers $j-k_i, k_i-k_{i-1},\cdots,k_2-k_1$. Applying the so-called "stars and bars" method in combinatorics, this is equal to

$$\binom{j-k_1+i-1}{i-1} \tag{20}$$

Thus,

$$
\begin{aligned}
2^j T_{i,j} &= \frac{1}{2^i} \sum_{k=1}^{j} \binom{j-k+i-1}{i-1} 2^k T_{0,k} \\
&= \frac{100}{2^i} \sum_{k=1}^{j} \binom{j-k+i-1}{i-1} 2^k \\
\Rightarrow T_{i,j} &= \frac{100}{2^{i+j}} \sum_{k=1}^{j} \binom{j-k+i-1}{i-1} 2^k
\end{aligned}
\tag{21}
$$

The average temperature of the red chunks after the final round round n is, therefore:

$$
\begin{aligned}
\overline{T_{\text{red}}} &= \frac{1}{n} \sum_{j=1}^{n} T_{n,j} \\
&= \frac{1}{n} \sum_{j=1}^{n} \frac{100}{2^{n+j}} \sum_{k=1}^{j} \binom{j-k+n-1}{n-1} 2^k \\
&= \frac{100}{n2^n} \sum_{j=1}^{n} \frac{1}{2^j} \sum_{k=1}^{j} \binom{j-k+n-1}{n-1} 2^k \\
&= \frac{100}{n2^n} \sum_{n \geq j \geq k \geq 1} \frac{1}{2^{j-k}} \binom{j-k+n-1}{n-1} \\
&= \frac{100}{n2^n} \sum_{j-k=0}^{n-1} \frac{n-(j-k)}{2^{j-k}} \binom{j-k+n-1}{n-1} \\
&= \frac{100}{n2^n} \sum_{d=0}^{n-1} \frac{n-d}{2^d} \binom{d+n-1}{n-1} \\
&= \frac{100}{n2^{2n}} \sum_{d=0}^{n-1} (n-d) 2^{n-d} \binom{2n-(n-d)-1}{n-1} \\
&= \frac{100}{n2^{2n}} \sum_{d=1}^{n} d2^d \binom{2n-d-1}{n-1} \\
&= \frac{100}{n2^{2n}} \cdot n \binom{2n}{n} \\
&= \frac{100}{2^{2n}} \binom{2n}{n}
\end{aligned}
\tag{22}
$$

The last step but two above uses a combinatorial identity:

$$
\sum_{k=1}^{n} k2^k \binom{2n-k-1}{n-1} = n \binom{2n}{n} \tag{23}
$$

of which a concise proof is given in the appendix section A.1.

Taking the limit when $n \to \infty$:

$$
\begin{aligned}
\lim_{n\to\infty} \overline{T_{\text{red}}} &= 100 \lim_{n\to\infty} \frac{1}{2^{2n}} \binom{2n}{n} \\
&= 100 \lim_{n\to\infty} \frac{\binom{2n}{n}}{\sum_{k=0}^{2n} \binom{2n}{k}} \\
&= 0
\end{aligned}
\tag{24}
$$

The last step above is still intuition-based, a proof of the limit is given in the appendix section A.2.

Finally, because of conservation of energy, we always have $\overline{T_{\text{red}}} + \overline{T_{\text{blue}}} = 100$, thus:

$$
\begin{aligned}
\lim_{n\to\infty} \Delta &= \lim_{n\to\infty} \left(\overline{T_{\text{blue}}} - \overline{T_{\text{red}}}\right) \\
&= \lim_{n\to\infty} \left(100 - 2\overline{T_{\text{red}}}\right) \\
&= 100 - 2 \lim_{n\to\infty} \overline{T_{\text{red}}} \\
&= 100
\end{aligned}
\tag{25}
$$

This completes the proof, and also the body of the solution. □

A Appendix

A.1 Proving the combinatorial identity

In this section I will prove the combinatorial identity I used above:

$$
\sum_{k=1}^{n} k2^k \binom{2n-k-1}{n-1} = n\binom{2n}{n} \tag{26}
$$

Proof. First, transform the left hand side:

$$
\begin{aligned}
\sum_{k=1}^{n} k2^k \binom{2n-d-1}{n-1} &= \sum_{k=1}^{n} k2^k \binom{2n-k-1}{n-k} \\
&= \sum_{k=1}^{\infty} k2^k \binom{2n-k-1}{n-k} \\
&= \sum_{k=0}^{\infty} (k+1)2^{k+1} \binom{2n-k-2}{n-k-1}
\end{aligned}
\tag{27}
$$

If we denote $t_k = (k+1)2^{k+1}\binom{2n-k-2}{n-k-1}$, then observe that:

$$
\begin{aligned}
\frac{t_{k+1}}{t_k} &= \frac{(k+2)2^{k+2}\binom{2n-k-3}{n-k-2}}{(k+1)2^{k+1}\binom{2n-k-2}{n-k-1}} \\
&= \frac{2(k+2)[k+(1-n)]}{(k+1)[k+(2-2n)]}
\end{aligned}
\tag{28}
$$

which is a rational expression with respect to k. Thus, the Gaussian hypergeometric function can be introduced:

$$\begin{aligned} \text{LHS} &= t_0 \cdot {}_2F_1\left(\begin{matrix} 2, 1-n \\ 2-2n \end{matrix}; 2\right) \\ &= 2\binom{2n-2}{n-1} {}_2F_1\left(\begin{matrix} 2, 1-n \\ 2-2n \end{matrix}; 2\right) \end{aligned} \tag{29}$$

Then the goal is simplify the hypergeometric term.

It will be a lot easier to evaluate ${}_2F_1\left(\begin{smallmatrix} a,b \\ c \end{smallmatrix}; z\right)$ when $z = 1$. To turn 2 into 1, apply Kummer's quadratic transformation:

$${}_2F_1\left(\begin{matrix} a, b \\ 2b \end{matrix}; z\right) = \frac{1}{(1-z)^{a/2}} {}_2F_1\left(\begin{matrix} \frac{1}{2}a, b - \frac{1}{2}a \\ b + \frac{1}{2} \end{matrix}; \frac{z^2}{4z-4}\right) \tag{30}$$

In our case, $a = 2, b = 1 - n, c = 2 - 2n$:

$${}_2F_1\left(\begin{matrix} 2, 1-n \\ 2-2n \end{matrix}; 2\right) = -{}_2F_1\left(\begin{matrix} 1, -n \\ \frac{3}{2}-n \end{matrix}; 1\right) \tag{31}$$

Since $\Re(1-n) < \Re\left(\frac{3}{2}-n\right)$, Gaussian's Hypergeometric Theorem applies:

$$\begin{aligned} {}_2F_1\left(\begin{matrix} 1, -n \\ \frac{3}{2}-n \end{matrix}; 1\right) &= \frac{\Gamma\left(\frac{3}{2}-n\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}-n\right)\Gamma\left(\frac{3}{2}\right)} \\ &= 1-2n \end{aligned} \tag{32}$$

Thus ${}_2F_1\left(\begin{smallmatrix} 2,1-n \\ 2-2n \end{smallmatrix}; 2\right) = 2n - 1$, and plugging it back:

$$\begin{aligned} \text{LHS} &= 2\binom{2n-2}{n-1}(2n-1) \\ &= 2n\binom{2n-1}{n-1} \\ &= n\binom{2n}{n} = \text{RHS} \end{aligned} \tag{33}$$

□

Note: I used hypergeometric functions simply because I happened to be reading about it and the hypergeometric approach is more straightforward. Elementary proofs can be found on `https://math.stackexchange.com/questions/3508257/proof-for-sum-k-1n-k2k-binom2n-k-1n-1-n-binom2nn/3508703`.

A.2 Proving the final limit

In this section I will prove that

$$\lim_{n\to\infty} 2^{-2n}\binom{2n}{n} = 0 \tag{34}$$

Proof. Apply Stirling's approximation:

$$\lim_{n\to\infty} \binom{2n}{n} \frac{\left[\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right]^2}{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}} = \lim_{n\to\infty} \left[\frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{n!}\right]^2 \frac{(2n)!}{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}} \tag{35}$$
$$= 1$$

Thus,

$$\begin{aligned}
\lim_{n\to\infty} 2^{-2n}\binom{2n}{n} &= \lim_{n\to\infty} 2^{-2n}\frac{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}}{\left[\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right]^2} \cdot \lim_{n\to\infty} \binom{2n}{n} \frac{\left[\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right]^2}{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}} \\
&= \lim_{n\to\infty} 2^{-2n}\frac{\sqrt{4\pi}2^{2n}n^{2n+\frac{1}{2}}e^{-2n}}{2\pi n^{2n+1}e^{-2n}} \qquad (36) \\
&= \lim_{n\to\infty} \frac{1}{\sqrt{\pi n}} \\
&= 0
\end{aligned}$$

□