Convex!

Solution of part (b) *of Problem* 99-002 *by* Ian Affleck (CECM, Simon Fraser University, Burnaby, BC, V5A 1S6 Canada (iaffleck@cs.sfu.ca)).

Claim. f_N is strictly convex for every $N \in \mathbb{N}$.

Proof. We show that for all $x, y \in \mathbb{R}_+^N$,

$$\frac{f_N(x) + f_N(y)}{2} > f_N\left(\frac{x+y}{2}\right). \tag{1}$$

First note that since $f_N(z) = cf_N(cz)$ for any $c \in \mathbb{R}_+$ and $z \in \mathbb{R}_+^N$, it suffices to show (1) for $x, y \in D^N$, where D^N is defined as

$$D^N := \left\{ z \in \mathbb{R}_+^N \ : \ \sum_{i=1}^{N} z_i \leq 1 \right\}.$$

Our approach is to express $f_N(z)$ for $z \in D^N$ as the expected time of a particular random variable concerning a coupon collecting problem.

One common form of the classic coupon collector problem sees an individual obtaining at most one of N types of coupons with each purchase made. Coupon types $C_1, C_2, \dots, C_N$ are distributed among the purchased items with respective proportions $z_1, z_2, \dots, z_N$ (so that the likelihood of obtaining *no* coupon with any given purchase is $1 - z_1 - z_2 - \cdots - z_N$). Of interest to us is the random variable $t(z)$ representing the minimum number of purchases required to obtain at least one coupon of each type, given $z = (z_1, z_2, \dots, z_N) \in D^N$. For ease of notation, we denote by $[n]$ the set $\{1, 2, \dots, n\}$. □

Lemma **1.** *Let $z \in D^N$ and let $k \in [N-1]$. The probability that C_{k+1} has been collected before the time that the last of $C_1, \dots, C_k$ is collected is*

$$\sum_{\emptyset \neq S \subseteq [k]} (-1)^{|S|+1} \frac{z_{k+1}}{z_{k+1} + \sum_{i \in S} z_i}.$$

Proof of Lemma 1. For $S \subseteq [k]$, let X_S be the the event that C_{k+1} is collected before the *first* of C_i, $i \in S$, has been collected. Then

$$P(X_S) = \frac{z_{k+1}}{z_{k+1} + \sum_{i \in S} z_i}. \tag{2}$$

Note C_{k+1} is collected before the *last* of $C_1, \dots, C_k$ if and only if $X_{\{i\}}$ occurs, for some $i \in [k]$. The probability of this union of events is

$$\sum_{1 \leq i \leq k} P(X_{\{i\}}) \ - \sum_{1 \leq i < j \leq k} P(X_{\{i\}} \cap X_{\{j\}}) \ + \ \cdots \ + \ (-1)^{k+1} P(X_{\{1\}} \cap X_{\{2\}} \cap \cdots \cap X_{\{k\}}),$$

by the principle of inclusion and exclusion. Since $P(X_S \cap X_T) = P(X_{S\cup T})$, we may rewrite this as

$$\sum_{\emptyset \neq S \subseteq [k]} (-1)^{|S|+1} P(X_S)\,.$$

By substituting (2), we establish Lemma 1. $\square$

The following lemma can be derived from a result in [2], but we include it for completeness.

LEMMA **2.** *Let $z \in D^N$. Then*

$$E[t(z)] = \sum_{\emptyset \neq S \subseteq [N]} (-1)^{|S|+1} \frac{1}{\sum_{i\in S} z_i} = f_N(z_1, \ldots, z_N)\,.$$

Proof of Lemma 2. The equality on the right (implicit in the statement of the problem) can be seen by expanding the product in f_N and integrating term by term. To prove the left equality, we proceed as follows. Let τ_k be the random variable representing the time at which the last of $C_1, \ldots, C_k$ is collected (with $\tau_0 = 0$), and define $\delta_k = \tau_k - \tau_{k-1}$ for $k \in [N]$, so that

$$E[t(z)] = E[\tau_N] = \sum_{i=1}^{N} E[\delta_i]\,.$$

We show by induction on m that

$$\sum_{i=1}^{m} E[\delta_i] = \sum_{\emptyset \neq S \subseteq [m]} (-1)^{|S|+1} \frac{1}{\sum_{i\in S} z_i}\,, \tag{3}$$

for $m \in [N]$. When $m = 1$, (3) reduces to the fact that the expectation of the geometric random variable δ_1 (time before C_1 is collected) is $\frac{1}{z_1}$. Assume (3) holds for $m = k \in [N-1]$. Under this induction hypothesis, it suffices to show that

$$E[\delta_{k+1}] = \sum_{\substack{S \subseteq [k+1] \\ k+1 \in S}} (-1)^{|S|+1} \frac{1}{\sum_{i\in S} z_i}\,.$$

Let ϕ_k be the event that C_{k+1} has been collected at time τ_k. Note $E[\delta_{k+1}|\phi_k] = 0$ and $E[\delta_{k+1}|\bar{\phi}_k] = \frac{1}{z_{k+1}}$, while $P(\phi_k)$ was found in Lemma 1. Then

$$\begin{aligned} E[\delta_{k+1}] &= E[\delta_{k+1} \mid \phi_k]\ P(\phi_k) + E[\delta_{k+1} \mid \bar{\phi}_k]\ P(\bar{\phi}_k) \\ &= 0 + \frac{1}{z_{k+1}} \left(1 - \sum_{\emptyset \neq S \in [k]} (-1)^{|S|+1} \frac{z_{k+1}}{z_{k+1} + \sum_{i\in S} z_i} \right) \end{aligned}$$

$$= \frac{1}{z_{k+1}} + \sum_{\emptyset \neq S \in [k]} (-1)^{|S|} \frac{1}{z_{k+1} + \sum_{i \in S} z_i}$$

$$= \sum_{k+1 \in S \subseteq [k+1]} (-1)^{|S|+1} \frac{1}{\sum_{i \in S} z_i},$$

completing the induction step, and Lemma 2 is established. □

Proof of Claim (continued). To prove (1), it suffices by Lemma 2 to show that for any two distinct vectors $x, y \in D^N$, we have

$$\frac{E[t(x)] + E[t(y)]}{2} > E\left[t\left(\frac{x+y}{2}\right)\right]. \tag{4}$$

To this end, we define a sequence $\{z^{(n)}\}_{n=1}^{\infty}$ of coupon distribution schemes, derived from x and y as follows: according to $z^{(n)}$, the coupon distribution during all of the first n purchases is constant, being either x or y with equal likelihood, while the coupon distribution on all subsequent purchases is $\frac{x+y}{2}$. As each $z^{(n)}$ describes a manner by which coupons are distributed, it induces a random variable $t(z^{(n)})$ similar to that of a vector in D^N. We note that

$$\lim_{n\to\infty} E[t(z^{(n)})] = \frac{E[t(x)] + E[t(y)]}{2} \qquad \text{and}$$

$$E[t(z^{(1)})] = E\left[t\left(\frac{x+y}{2}\right)\right].$$

We aim to show that $E[t(z^{(n)})] < E[t(z^{(n+1)})]$ for all $n \geq 1$.

Fix $n \geq 1$, and for ease of notation, let $z = z^{(n)}$ and $z' = z^{(n+1)}$. For positive integer k and $i \in [N]$, let $m_i(k)$ be the probability that coupon type C_i has not been obtained after k purchases using z, and define $m_i'(k)$ similarly for z'. Then

$$m_i(n) = m_i'(n) = \frac{(1-x_i)^n + (1-y_i)^n}{2}.$$

Meanwhile,

$$m_i'(n+1) = \frac{1}{2}((1-x_i)^{n+1} + (1-y_i)^{n+1}) \qquad \text{and}$$

$$m_i(n+1) = \frac{1}{4}((1-x_i)^{n+1} + (1-x_i)^n(1-y_i) + (1-y_i)^n(1-x_i) + (1-y_i)^{n+1}).$$

Thus,

$$\begin{aligned}
&4(m_i'(n+1) - m_i(n+1)) \\
&\quad = (1-x_i)^{n+1} - (1-x_i)^n(1-y_i) - (1-y_i)^n(1-x_i) + (1-y_i)^{n+1} \\
&\quad = ((1-x_i)^n - (1-y_i)^n)((1-x_i) - (1-y_i)) \;\geq\; 0,
\end{aligned} \tag{5}$$

with strict inequality if and only if $x_i \neq y_i$.

Let ρ_i be the probability that coupon type C_i is collected for the first time on the $(n+1)$th purchase using z, and define ρ'_i similarly for z'. Then by (5),

$$\rho_i = m_i(n) - m_i(n+1) \geq m'_i(n) - m'_i(n+1) = \rho'_i\,. \tag{6}$$

That is, for $i \in [N]$, coupon type C_i is at least as likely to be collected on the $(n+1)$th purchase using z as opposed to using z'. We claim that this fact implies $E[t(z)] < E[t(z')]$.

Let $C = \{C_1, \ldots, C_N\}$. For each $W \subseteq C$ and $k \geq 1$, define $P_W(k)$ (resp., $P'_W(k)$) to be the probability that W is the set of all coupon types which have been collected after k purchases using z (resp., z'), and define E_W to be the expected number of purchases required to collect all coupons in $C - W$ when coupons are distributed according to $\frac{x+y}{2}$. In the following expansion, we have interpreted W to be the set of coupon types that have been collected by the end of the nth purchase:

$$\begin{aligned} E[t(z)] &= \sum_{j=1}^{n} j(P_C(j) - P_C(j-1)) \\ &\quad + (1 - P_C(n))\left((n+1) + \sum_{W \subseteq C} P_W(n)\left(\sum_{C_i \notin W} \rho_i E_{W \cup \{C_i\}} + \left(1 - \sum_{C_i \notin W} \rho_i\right) E_W\right)\right). \end{aligned}$$

Similarly,

$$\begin{aligned} E[t(z')] &= \sum_{j=1}^{n} j(P'_C(j) - P'_C(j-1)) \\ &\quad + (1 - P'_C(n))\left(n+1 + \sum_{W \subseteq C} P'_W(n)\left(\sum_{C_i \notin W} \rho'_i E_{W \cup \{C_i\}} + \left(1 - \sum_{C_i \notin W} \rho'_i\right) E_W\right)\right). \end{aligned}$$

Because z and z' are identical for the first n purchases, we have $P_C(j) = P'_C(j)$, for all $j \in [n]$, and $P_W(n) = P'_W(n)$ for every $W \subseteq C$. Thus

$$E[t(z)] - E[t(z')] = (1 - P_C(n))\left(\sum_{W \subseteq C} P_W(n)\left(\sum_{C_i \notin W} (\rho_i - \rho'_i)(E_{W \cup \{C_i\}} - E_W)\right)\right).$$

From (6), we have that $\rho_i - \rho'_i \geq 0$, for all i. Since x and y are distinct, this inequality is strict for some i. Furthermore, whenever $C_i \notin W$ we have $E_W > E_{W \cup \{C_i\}}$. Thus

$$E[t(z)] - E[t(z')] < 0.$$

As this inequality holds for all $n \geq 1$, we have

$$E\left[t\left(\frac{x+y}{2}\right)\right] = E[t(z^{(1)})] < \lim_{n \to \infty} E[t(z^{(n)})] = \frac{E[t(x)] + E[t(y)]}{2},$$

for all distinct $x, y \in D^N$, as desired. □

Remarks on the Solution. Our study of the coupon collector problem originated as an investigation of expected time to broadcast in *unreliable* arborescence networks using randomized strategies. (An unreliable network is one in which each edge fails to transmit the broadcast message with fixed probability on each attempt, and the success or failure of each attempt is unknown to the caller.) A randomized broadcast strategy (RBS) assigns to each parent vertex a probability distribution on its out-neighbor set, by which it chooses a neighbor to call at each time step after receiving the message. Verification of the above claim implies that each unreliable star network has a unique optimal RBS, when optimality is measured with respect to expected time to broadcast from the center. In [1], we extend the above proof to show that the optimal RBS is unique for every unreliable arborescence.

A more direct analytic proof of the strict convexity of $f_N(x)$ has recently been proposed by Borwein and Hijab (see [2]). They observe that $f_N(x)$ can be expressed as an expectation

$$\mathrm{E}\left[\max\left(\frac{X_1}{x_1}, \frac{X_2}{x_2}, \ldots, \frac{X_N}{x_N}\right)\right] ,$$

where the X_i are independent positive-valued random variables satisfying

$$\mathrm{P}[X_i > t] = e^{-t} ,$$

for $i \in [N]$, and obtain the desired result using a Laplace transform. This method of proof in fact shows concavity of $1/f_N$, part (d) of the submitted problem.

The author wishes to thank Luis Goddyn and Jonathan Borwein for their helpful comments, suggestions and discussions.

References

[1] I. A. Affleck, *Minimizing Expected Broadcast Time in Unreliable Networks*, Ph.D. thesis, Simon Fraser University, 2000.

[2] J. Borwein and O. Hijab, *Solution to problem 99-002: Convexity?*, personal communication (2000).

[3] H. B. Nath, *Waiting time in the coupon collecter problem*, Austral. N.Z. J. Stat., 15 (1973), pp. 132 - 135.