MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Spring 2005)

Problem Set 2G: Solutions
Due: February 16, 2005

G1. Let us compute the probability of selecting a first-born child using the two methods. Denote by E the event of interest (selecting a first-born child).

(i) Let A_i be the event that the selected family has i children. Using the Total Probability Theorem and the multiplication rule,

$$\mathbf{P}(E) = \sum_{i=1}^{k} \mathbf{P}(E \cap A_i) = \sum_{i=1}^{k} \mathbf{P}(A_i)\mathbf{P}(E|A_i) \quad .$$

The family is first selected uniformly; hence, the probability of selecting any one of them is $1/m$ and because there are n_i with i children, it follows that

$$\mathbf{P}(A_i) = \frac{n_i}{m} = \frac{n_i}{\sum_{i=1}^{k} n_i} \quad .$$

Given that a family with i children has been selected, the probability of selecting at random (uniformly) the first-born child is

$$\mathbf{P}(E|A_i) = \frac{1}{i} \quad .$$

Putting it all together, we have

$$\mathbf{P}(E) = \left(\sum_{i=1}^{k} \frac{n_i}{i}\right)\frac{1}{m} = \left(\sum_{i=1}^{k} \frac{n_i}{i}\right)\frac{1}{\sum_{i=1}^{k} n_i} \tag{1}$$

(ii) Using this method a child is selected directly in a uniform fashion. Because there are $\sum_{i=1}^{k} i n_i$ children and m first-born ones,

$$\mathbf{P}(E) = \frac{m}{\sum_{i=1}^{k} i n_i} \tag{2}$$

Now we compare expressions (1) and (2) and, to prove the proposition, we need to establish that

$$\left(\sum_{i=1}^{k} i n_i\right)\left(\sum_{j=1}^{k} \frac{n_j}{j}\right) \geq \left(\sum_{i=1}^{k} n_i\right)\left(\sum_{j=1}^{k} n_j\right) \quad .$$

Expanding the left-hand side,

$$\begin{aligned}
\left(\sum_i i n_i\right)\left(\sum_j \frac{n_j}{j}\right) &= \sum_i \sum_j \left(\frac{i}{j}\right) n_i n_j = \sum_i \left(n_i^2 + \sum_{i \neq j} \left(\frac{i}{j}\right) n_i n_j\right) \\
&= \sum_i \left(n_i^2 + \sum_{j<i} \left(\frac{i}{j}\right) n_i n_j + \sum_{j>i} \left(\frac{i}{j}\right) n_i n_j\right) \\
&= \sum_i n_i^2 + \sum_i \sum_{j<i} \left(\frac{i}{j}\right) n_i n_j + \sum_i \sum_{j>i} \left(\frac{i}{j}\right) n_i n_j
\end{aligned}$$

$$\begin{aligned}
&= \sum_i n_i^2 + \sum_j \sum_{i>j} \left(\frac{i}{j}\right) n_i n_j + \sum_i \sum_{j>i} \left(\frac{i}{j}\right) n_i n_j \\
&= \sum_i n_i^2 + \sum_i \sum_{j>i} \left(\frac{j}{i}\right) n_j n_i + \sum_i \sum_{j>i} \left(\frac{i}{j}\right) n_i n_j \\
&= \sum_i n_i^2 + \sum_i \sum_{j>i} \left(\frac{i}{j} + \frac{j}{i}\right) n_i n_j = \sum_i \left(n_i^2 + \sum_{j>i} \left(\frac{i^2+j^2}{ij}\right) n_i n_j \right)
\end{aligned}$$

Expanding the right-hand side,

$$\left(\sum_i n_i\right)\left(\sum_j n_j\right) = \sum_i \sum_j n_i n_j = \sum_i \left(n_i^2 + \sum_{j \neq i} n_i n_j \right) = \sum_i \left(n_i^2 + 2\sum_{j>i} n_i n_j \right)$$

Comparing like terms in the expansion of each side, the result is established provided

$$\frac{i^2+j^2}{ij} \geq 2 \quad ,$$

which is indeed satisfied given $i, j = 1, 2, \ldots k$.

G2. (a) Let U_F be the event that the connection F to B is unblocked, so $\mathbf{P}(U_F) = 0.85$, $\mathbf{P}(U_F^c) = 0.15$. Let U_A be the event that the connection A to B is unblocked (i.e., at least one of the paths from A to B is unblocked). We have

$$\begin{aligned}
\mathbf{P}(U_C \mid U_F) &= 1-(1-0.8\cdot 0.9)(1-0.95) = 0.986 \\
\mathbf{P}(U_C \mid U_F^c) &= 0.8 \cdot 0.9 \\
\mathbf{P}(U_D \mid U_F) &= 1-(1-0.95)(1-0.9) = 0.995 \\
\mathbf{P}(U_D \mid U_F^c) &= 0.95 \\
\mathbf{P}(U_A \mid U_F) &= 1-(1-0.9\cdot \mathbf{P}(U_C \mid U_F))(1-0.75\cdot \mathbf{P}(U_D \mid U_F)) = 0.9714 \\
\mathbf{P}(U_A \mid U_F^c) &= 1-(1-0.9\cdot \mathbf{P}(U_C \mid U_F^c))(1-0.75\cdot \mathbf{P}(U_D \mid U_F^c)) = 0.8988 \\
\mathbf{P}(U_A) &= \mathbf{P}(U_F)\mathbf{P}(U_A \mid U_F) + \mathbf{P}(U_F^c)\mathbf{P}(U_A \mid U_F^c) = \boxed{0.9605}.
\end{aligned}$$

(b) The line of analysis of part (a) applies. Using the probability $\mathbf{P}(U_A)$ calculated in part (a), the desired probability is

$$1-(1-\mathbf{P}(U_A))(1-0.8) = 1-(1-0.9605)(1-0.8) = \boxed{0.9921}.$$

G3. Let $p_{i,n-i}(k)$ denote the probability that after k exchanges, a jar will contain i balls that started in that jar and $n-i$ bars that started in the other jar. We want to find $p_{n,0}(4)$. We argue recursively, using the total probability theorem. We have

$$\begin{aligned}
p_{n,0}(4) &= \frac{1}{n} \cdot \frac{1}{n} \cdot p_{n-1,1}(3), \\
p_{n-1,1}(3) &= p_{n,0}(2) + 2 \cdot \frac{n-1}{n} \cdot \frac{1}{n} \cdot p_{n-1,1}(2) + \frac{2}{n} \cdot \frac{2}{n} \cdot p_{n-2,2}(2),
\end{aligned}$$

$$\begin{aligned}
p_{n,0}(2) &= \frac{1}{n} \cdot \frac{1}{n} \cdot p_{n-1,1}(1), \\
p_{n-1,1}(2) &= 2 \cdot \frac{n-1}{n} \cdot \frac{1}{n} \cdot p_{n-1,1}(1) \\
p_{n-2,2}(2) &= \frac{n-1}{n} \cdot \frac{n-1}{n} \cdot p_{n-1,1}(1), \\
p_{n-1,1}(1) &= 1.
\end{aligned}$$

Combining these equations, we obtain

$$p_{n,0}(4) = \frac{1}{n^2}\left(\frac{1}{n^2} + \frac{4(n-1)^2}{n^4} + \frac{4(n-1)^2}{n^4}\right).$$