

1

1.1

For any fixed input x , every time we run our randomized algorithm \bar{A} , the performance of \bar{A} on x is determined. Say our algorithm \bar{A} on given input x has a probability of p_i to perform as deterministic algorithm A_i , then since our given input x is definite and finite, we can always find a set $S = \{A_1, A_2, \dots, A_m\}$ and some corresponding possibilities p_1, p_2, \dots, p_m where $\sum_{i=1}^m p_i = 1$, by which the performance of our algorithm \bar{A} belongs to S . So the expected running time of \bar{A} on input x is

$$\mathbb{E} [T(\bar{A}, x)] = \sum_{i=1}^m p_i T(A_i, x)$$

Then we define distribution \mathcal{A} to be

$$A = A_i \quad \text{with probability} \quad p_i, \forall i \in [m]$$

Clearly

$$\mathbb{E}_{A \sim \mathcal{A}} [T(A, x)] = \sum_{i=1}^m p_i T(A_i, x)$$

So for any randomized algorithm \bar{A} we can find a distribution \mathcal{A} over \mathcal{A} to represent the expected running time of \bar{A}

As for the expected running time of the best algorithm for Π , since it is the best algorithm, it must have the minimum expected running time, i.e. the $\mathbb{E}_{A \sim \mathcal{A}} [T(A, x)]$ for this algorithm must be the smallest. We also want the expected running time to be upper bounded by the worst input distribution, which means that we would like to minimize the running time over the maximum among $x \in \mathcal{X}$. So the overall running time of the best randomized algorithm can be optimized by finding the best distribution \mathcal{A} over \mathcal{A} and the best randomized algorithm should follow the distribution of \mathcal{A} . So the expected running time of the best algorithm for Π is

$$\min_{\mathcal{A} \sim \mathcal{A}} \max_{x \in \mathcal{X}} \mathbb{E}_{A \sim \mathcal{A}} [T(A, x)]$$

1.2

To use Von Neumann's minimax theorem, we would like to show that the expected running time can be viewed as elements in a matrix.

Imagine that we have two players denoted by the Column and Row players. Our Column player has a set of algorithms to solve a question and our Row player has the set of all

possible inputs for such a question. Denote the running time of specific algorithm a on given input x to be $T(a, x)$. The goal of Column player is to optimize his distribution \mathcal{A} so that for any distribution of inputs given by Row player, the expected running time is smallest. While the goal of Row player is to optimize his distribution \mathcal{X} so that for any distribution of algorithms given by Column player, the expected running time is largest. Any element of the matrix can be expressed as $T(a, x)$ where $T(a, x)$ is the running time of algorithm a on input x . So the goal of the Column player can be expressed as

$$\min_{\mathcal{A} \sim \mathcal{A}} \max_{x \in \mathcal{X}} \mathbb{E}_{A \sim \mathcal{A}} [T(A, x)]$$

And the goal of the Row player can be expressed as

$$\max_{\mathcal{X} \sim \mathcal{X}} \min_{a \in \mathcal{A}} \mathbb{E}_{X \sim \mathcal{X}} [T(a, X)]$$

Then Von Neumann's minimax theorem says there exists mixed strategies that are optimal for both players and this mixed strategies provide the same optimal value for them, which means that

$$\max_{\mathcal{X} \sim \mathcal{X}} \min_{a \in \mathcal{A}} \mathbb{E}_{X \sim \mathcal{X}} [T(a, X)] = \min_{\mathcal{A} \sim \mathcal{A}} \max_{x \in \mathcal{X}} \mathbb{E}_{A \sim \mathcal{A}} [T(A, x)]$$

1.3

1.4

2

2.1

We first construct a weighted directed graph H from G . Then we can find maxflow on graph H to see whether there is a perfect matching in G .

Firstly, we let all the edges in G point from V_1 to V_2 in H , and we assign them a weight of $+\infty$.

Then we add a source vertex s and add an edge with weight 1 from s to all vertices in V_1 .

Finally we add a sink vertex t and add an edge with weight 1 from all vertices in V_2 to t .

Now we have constructed the graph H , we can run the Network Flow algorithm to find a maxflow of H . If the maxflow of H , denoted by f , turns out to be $f = |V_1| = |V_2|$, then there exists a perfect matching in G . And the perfect matching contains exactly the edges in G used by us to construct the maxflow of H .

2.2

Proof of Necessity:

This part is very straightforward.

Suppose there is a perfect matching M from V_1 to V_2 . Then for any $S \subset V_1$, for every vertex $v \in S$, there is an edge in M connecting v to a vertex in V_2 . This means that there are at least as many vertices in V_2 that are neighbors of vertices in V_1 as there are vertices in V_1 .

That is to say, for any $S \subset V_1$, $|N(S)| \geq |S|$.

Proof of Sufficiency:

Following the hint, we would like to prove by showing that the mincut of the graph we construct is exactly $|V_1|$ (or $|V_2|$ if you like).

Firstly, any mincut can only contain the edges in H which are not in G because we assign edges in G with a weight of $+\infty$.

Then there is a cut with capacity $|V_1|$ if we make our S - T cut to be $S = \{s\}$, the singleton. So the capacity of mincut of H is at most $|V_1|$.

Suppose there is another minimum S - T cut where $S \setminus V_2 = \{s\} \cup (V_1 \setminus V'_1)$ and $T \setminus V_1 = \{t\} \cup (V_2 \setminus V'_2)$, since this is a mincut, there are no edges from $V_1 \setminus V'_1$ to $V_2 \setminus V'_2$. This means that all the neighbors of $V_1 \setminus V'_1$ must be in V'_2 . Then by the property that $|N(S)| \geq |S|$ for any $S \subset V_1$, $|V_1 \setminus V'_1| \leq |V'_2|$.

Finally the capacity of this cut is $|V'_1| + |V'_2|$, $|V'_1| + |V'_2| \geq |V'_1| + |V_1 \setminus V'_1| = |V_1|$. This means that the capacity of any S - T cut is at least $|V_1|$. Then by our instance where $S = \{s\}$, the mincut is indeed $|V_1|$. Finally by the Maxflow-Mincut Theorem, the maxflow of H equals its mincut $|V_1|$, which means that there is a perfect matching in G .

3

We prove by showing that the debt network G can be equivalently transformed to another network H which has $n - 1$ edges at most.

In the original debt graph G , for any vertex $v \in V$, we define

$$W(v) = \sum_{(u,v) \in E} w(u,v) - \sum_{(v,p) \in E} w(v,p)$$

$W(v)$ is the amount of money that v owes other roommates or other roommates owe v . We then divide V into a partition $V = V_1 \cup V_2 \cup V_3$, where $\forall v \in V_1, W(v) > 0, \forall v \in V_2, W(v) < 0, \forall v \in V_3, W(v) = 0$.

To sum up, V_1 is the set of all creditors, V_2 is the set of all debtors and V_3 is the set of all people who do not need to pay or receive money from anyone else. In graph $H = (V', E')$ we do not have to consider vertices in V_3 , so $V' = V_1 \cup V_2$.

Then we pick anyone from V_2 and call him the pivot p . We construct the graph H by the following method:

For all vertices v in V_1 , we add an edge (p, v) with weight $W(v)$.

For all vertices v in $V_2 \setminus \{p\}$, we add an edge (v, p) with weight $W(v)$.

I.e., all debtors excluding pivot p give money to p , then p gives money to all creditors (on behalf of all debtors). So the total times it takes to carry out person-to-person payments is at most $|V_1| + |V_2| - 1$, which is at most $n - 1$.

Proof of Correctness:

The reason why H and G are equivalent when deciding the amount of payments is that graph H guarantees that every single person is receiving or giving out exactly the amount of money that he should collect or in debt. So our construction of H can indeed make sure that the debts are paid off at last.

4 Comments

4.1

4.2

4.3