Warsaw PhD Open Course: From Joins to Aggregates and Optimization Problems - Exam solutions

Djordje Zivanovic

January 26, 2019

1 Let $Q(\mathbf{A_1} \cup \mathbf{A_2} \cup \cdots \cup \mathbf{A_n})$ denote a join query on relations $R_1(\mathbf{A_1}) = R_1(A_1, A_2), \cdots R_n(\mathbf{A_n}) = R_n(A_n, A_1)$. Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ denote a corresponding hypegraph to the join query, then $\mathcal{V} = \{A_1, .., A_n\}$ and $\mathcal{E} = \{\{A_1, A_2\}, ..., \{A_n, A_1\}\}$.

1.1 Let x_{R_i} denote the weight of an hyperedge (relation) R_i in the join query Q. In addition let $cover_i$ denote the set of relations corresponding to edges that are part of the cover for the variable i. The fractional edge cover number $\rho^*(Q)$ is the cost of an optimal solution to the linear program:

$$\text{minimize} \sum_{i \in [n]} x_{R_i} \tag{1}$$

$$\text{subject to} \sum_{rel_j \in cover_i} x_{rel_j} \geq 1, i \in [n] \tag{2}$$

$$x_{R_i} \geq 0, i \in [n]$$

By a definition of a cover, only edges incident to variable are part of its cover. Thus, in our case $cover_i = \{R_{i-1}, R_i\}$, where R_0 is R_n in the case of A_1. If we rewrite the inequalities 2 we will get:

$$\begin{aligned} x_{R_1} + x_{R_n} &\geq 1 \\ x_{R_1} + x_{R_2} &\geq 1 \\ &\vdots \\ x_{R_{n-1}} + x_{R_n} &\geq 1 \end{aligned}$$

By summing inequalities we get:

$$2\left(\sum_{i \in [n]} x_{R_i}\right) \geq n$$

This means the minimized function $\rho^*(Q)$ cannot be less than $\frac{n}{2}$. This is achievable if we set $x_{R_i} = \frac{1}{2}, i \in [n]$. We can see that this solution satisfies all constraints in the linear program. Finally, $\rho^*(Q) = \frac{n}{2}$.

1.2 sdf