

Topics in Database Theory – Homework 3

1 The AGM Bound

1. (0 points)

(a) Consider the following query:

$$Q(x, y, z, u, v, w) = R(x, y) \wedge S(y, z) \wedge T(y, u) \wedge K(u, v) \wedge M(x, w)$$

Assume that $|R| = |S| = |T| = |K| = |M| \leq N$.

- i. Find the maximum size of the output to the query Q
- ii. Find a worst-case database instance where the query Q has the bound you found above.

(b) Consider the query:

$$Q(x, y, z, u) = R(x, y) \wedge S(y, z) \wedge T(z, u) \wedge K(u, x)$$

Suppose the four relations have cardinalities N_1, N_2, N_3, N_4 .

Give a formula that represents a tight upper bound on $|Q|$. Your formula should use the cardinalities N_1, N_2, N_3, N_4 and operations like $+$, \times , $/$, \wedge , \max , for example $\max(N_1/N_2, N_3^{3/2} + N_4)$ (not a real answer).

- (c) Consider the same query as above, and repeat your answer for the case when y is a key in S :

$$Q(x, y, z, u) = R(x, y) \wedge S(\underline{y}, z) \wedge T(z, u) \wedge K(u, x)$$

2 Information Inequalities

2. (0 points)

- (a) Let Y, Z be two finites sets. Prove that

$$|Y|^2 + |Z|^2 \leq |Y \cup Z|^2 + |Y \cap Z|^2$$

and that equality holds iff $Y \subseteq Z$ or $Z \subseteq Y$. (We used this property when we proved the generalized Shearer inequality.)

- (b) Consider the following query:

$$Q(x, y, z, u) = R(x, y, z) \wedge S(y, z, u) \wedge T(z, u, x) \wedge K(u, x, y)$$

Prove that the following inequalities hold:

$$|Q| \leq (|R| \cdot |S| \cdot |T| \cdot |K|)^{1/3}$$

$$|Q| \leq |R| \cdot \max(\deg_S(u|yz))$$

$$|Q| \leq |T| \cdot \max(\deg_K(y|ux))$$

- (c) Consider the following query:

$$Q(x, y, z, u, v, w) = R(x, y, z) \wedge S(z, u, v) \wedge T(v, w, x) \\ \wedge A(y, z, u) \wedge B(u, v, w) \wedge C(w, x, y)$$

Prove the following inequality:

$$|Q| \leq \sqrt{|R| \cdot |S| \cdot |T| \cdot \max(\deg_A(y|zu)) \cdot \max(\deg_B(u|vw)) \cdot \max(\deg_C(w|xy))}$$

- (d) Prove the following inequality:

$$h(xyz) + h(zuv) + h(vwx) + h(yuw) + \\ h(y|x) + h(z|y) + h(u|z) + h(v|u) + h(w|v) + h(x|w) \geq 3h(xyzuvw)$$

More details about information inequalities can be found in [1].

References

- [1] D. Suci. Applications of information inequalities to database theory problems. In *LICS*, pages 1–30, 2023.