

Описание Менеджера эвристик

Из сложности DPsize [Moe]:

$$I_{DPsize}^{chain}(n) = \begin{cases} \frac{1}{48}(5n^4 + 6n^3 - 14n^2 - 12n), & n \text{ even} \\ \frac{1}{48}(5n^4 + 6n^3 - 14n^2 - 6n + 11), & n \text{ odd} \end{cases}$$
$$I_{DPsize}^{cycle}(n) = \begin{cases} \frac{1}{4}(n^4 - n^3 - n^2), & n \text{ even} \\ \frac{1}{4}(n^4 - n^3 - n^2 + n), & n \text{ odd} \end{cases}$$
$$I_{DPsize}^{star}(n) = \begin{cases} 2^{2n-4} - \frac{1}{4}\binom{2n}{n-1} + q(n), & n \text{ even} \\ 2^{2n-4} - \frac{1}{4}\binom{2(n-1)}{n-1} + \frac{1}{4}\binom{n-1}{(n-1)/2} + q(n), & n \text{ odd} \end{cases}$$
$$\text{with } q(n) = n2^{2n-1} - 5 \times 2^{n-3} + \frac{1}{2}(2^n - 5n + 4)$$
$$I_{DPsize}^{clique}(n) = \begin{cases} 2^{2n-2} - 5 \times 2^{n-2} + \frac{1}{4}\binom{2n}{n} - \frac{1}{4}\binom{n}{n/2} + 1, & n \text{ even} \\ 2^{2n-2} - 5 \times 2^{n-2} + \frac{1}{4}\binom{2n}{n} + 1, & n \text{ odd} \end{cases}$$

Заметим, что полные графы clique обладают высокой алгоритмической сложностью планирования алгоритмом DPsize. Поэтому все полные графы от 4 и будем отдавать Geoq. Заметим, что также звёзды обладают высокой сложностью. Их можно разбить на цепи по центральной вершине. Посмотрим на сложность планирования цепей и циклов: при количестве таблиц ≤ 12 сложность циклов не более чем в 2 раза больше сложности цепи. Таким образом, при количестве таблиц ≤ 12 в цепи(цикле) будем применять стандартный DP, иначе Geoq.

Пусть дан большой (≥ 50 таблиц) аналитический запрос с внутренними соединениями. Для его планирования будем выполнять следующие шаги:

1. Представим запрос в виде связного графа, иначе выполним шаги 2 - 4 для каждой связной компоненты графа и соединим их декартовым произведением.
2. Разобьём граф топологии: сначала выделим clique (с 4+ вершинами), затем циклы, найдём звёзды и разобьём их на цепи.
3. Спланируем каждую топологию графа.
4. Алгоритмом **GOO** [Neu, стр. 108] [All18] соберём все планы топологий в один план.

Список литературы

- [All18] Julian Reddy Allam. Evaluation of a greedy join-order optimization approach using the imdb dataset. *University of Magdeburg*, page 23, 2018.
- [Moe] Guido Moerkotte. Dynamic programming for join ordering revisited. *School of Electrical Engineering and Computer Science, Faculty of Engineering University of Ottawa*, page 10.
- [Neu] Thomas Neumann. Query optimization. *Technical University of Munich*.