数据库原理CH15作业

18340040 冯大纬

15.1

Assume (for simplicity in this exercise) that only one tuple fits in a block and memory holds at most three blocks. Show the runs created on each pass of the sort-merge algorithm when applied to sort the following tuples on the first attribute: (kangaroo, 17), (wallaby, 21), (emu, 1), (wombat, 13), (platypus, 3), (lion, 8), (warthog, 4), (zebra, 11), (meerkat, 6), (hyena, 9), (hornbill, 2), (baboon, 12).

Answer:

第一轮：

- (emu,1), (kangaroo, 17), (wallaby, 21)
- (lion, 8), (platypus, 3), (wombat, 13)
- (meerkat, 6), (warthog, 4), (zebra, 11)
- (baboon, 12), (hornbill, 2), (hyena, 9)

第二轮，归并前三路：

- (emu,1), (kangaroo, 17), (lion, 8), (meerkat, 6), (platypus, 3), (wallaby, 21), (warthog, 4), (wombat, 13), (zebra, 11)
- (baboon, 12), (hornbill, 2), (hyena, 9)

第三轮，归并两路

- (baboon, 12), (emu,1), (hornbill, 2), (hyena, 9), (kangaroo, 17), (lion, 8), (meerkat, 6), (platypus, 3), (wallaby, 21), (warthog, 4), (wombat, 13), (zebra, 11)

15.3

Let relations r1(A, B, C) and r2(C, D, E) have the following properties: r1 has 20,000 tuples, r2 has 45,000 tuples, 25 tuples of r1 fit on one block, and 30 tuples of r2 fit on one block. Estimate the number of block transfers and seeks required using each of the following join strategies for r1 ⋈ r2:

a. Nested-loop join.

b. Block nested-loop join.

c. Merge join.

d. Hash join.

Answer:

Number of records of r1: n_{r1} = 20000

Number of blocks of r1: b_{r1} = 800

Number of records of r2: n_{r2} = 45000

Number of blocks of r2: b_{r2} = 1500

a. Nested-loop join.

$$\begin{aligned} total_{r1,r2} &= n_{r1} \times b_{r2} + b_{r1} = 20000 \times 1500 + 800 = 30,000,800 \\ total_{r2,r1} &= n_{r2} \times b_{r1} + b_{r2} = 45000 \times 800 + 1500 = 36,001,500 \end{aligned} \tag{1}$$

b. Block nested-loop join.

$$\begin{aligned} total_{r1,r2} &= b_{r1} \times b_{r2} + b_{r1} = 800 \times 1500 + 800 = 1,200,800 \\ total_{r2,r1} &= b_{r2} \times b_{r1} + b_{r2} = 1500 \times 800 + 1500 = 1,201,500 \end{aligned} \tag{2}$$

c. Merge join.

假设内存容量一次可以读入M个block

$$total = 1500(2\lceil log_{M-1}(1500/M) \rceil + 2) + 800(2\lceil log_{M-1}(800/M) \rceil + 2) + 1500 + 800 \tag{3}$$

d. Hash join.

$$total = 2(1500 + 800)\lceil log_{M-1}(800) - 1 \rceil + 1500 + 800 \tag{4}$$