Database Systems, CSCI 4380-01
Example

1 Exercise

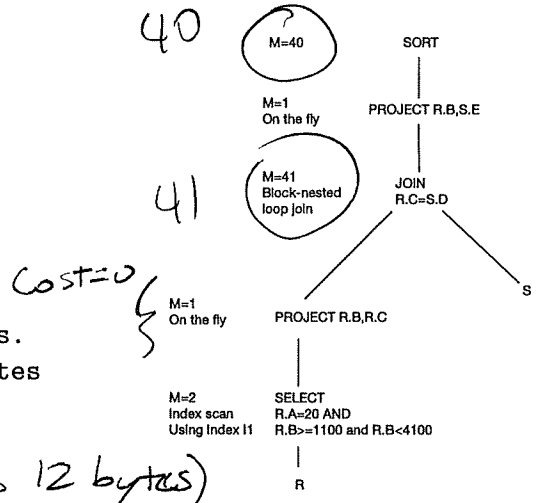
You are given the following information and statistics. Find the expected cost of the query plan below.

```
R(A,B,C,K,L,M)  10,000,000 tuples
    Distinct(R.A)= 200
    Distinct(R.B)=5,000
          MinVal(R.B)=500    MaxVal(R.B)=10,500
    Distinct(R.C)=6,000
S(C,D,E,F,G,H)  100,000 tuples
    Distinct(S.D)=5,000

Each attribute in the given relations is 12 bytes.
A disk page and a memory block both have 8000 bytes
       of space for data.

Index I1 is on R(A,B,C).
```

(each pointer is 12 bytes)



Fill in the information below to help solve this problem:

1. How many tuples fit in a page for a relation with 2 attributes, 3 attributes and 6 attributes?

2 attr → $2 \times 12 = 24$ bytes → $\left\lfloor \frac{8000}{24} \right\rfloor = 333$ tuples/page

3 attr → 36 bytes

6 attr → 72 bytes → $\left\lceil \frac{8000}{36} \right\rceil = 222$ //

→ 111 //

2. How many pages are in relations R and S?

$\frac{10\,000\,000}{111} \cong 90{,}000$ pages

$\frac{100{,}000}{111} \cong 900$ pages

3. How many leaf nodes and levels are in index I1?

Each tuple $12 + 3 \times 12 = 48$ (12 ↑ pointer)

$\left\lfloor \frac{8000}{48} \right\rfloor = 166$ tuples/node

leaf $\left\lceil \frac{10{,}000{,}000}{166} \right\rceil = 60241$ leaf nodes

level 1 $\left\lceil \frac{60241}{166} \right\rceil = 363$

4 level → 60241 leaf nodes

level $\left\lceil \frac{363}{166} \right\rceil = 3$ + root

4. How many tuples are expected after the selection: SELECT_{R.A=20 AND R.B>=1100 AND R.B<4100?

R.A=20 & R.B≥1100 & R.B≤4100

$\underbrace{\frac{1}{200} * \frac{3000}{10000}}_{sel(c)} * 1000000 = 15{,}000$ tuples

5. What is the size of relation after projection over the selection above (in terms of number of pages)?

$\frac{15000}{333} \simeq 45$ pages

6. What is the cost of the block nested loop join?

COST OF READING R = 0 [Already in memory from index search]

READ S TWO TIMES = 900 * 2 = 1,800 pages

7. How many tuples are in the result of join operation (including selection)?

$\underbrace{\frac{1}{200} * \frac{3000}{10000} * \frac{1}{6000}}_{sel} * 10{,}000{,}000 * 100{,}000 =$ 250,000 tuples

8. What is the size of relation after projection over the join (in terms of number of pages)?

$\frac{250{,}000}{333} =$ 750 pages

9. What is the cost of the sort? (M=40)

750 (write once) / $\lceil \frac{750}{40} \rceil = 19$ sorted groups

750 (read + merge & output)

= 1500 pages

10. What is the total expected cost of this query plan?

94 + 1,800 + 1,500 = 3,394 pages

3(a). What is the cost of index scan?

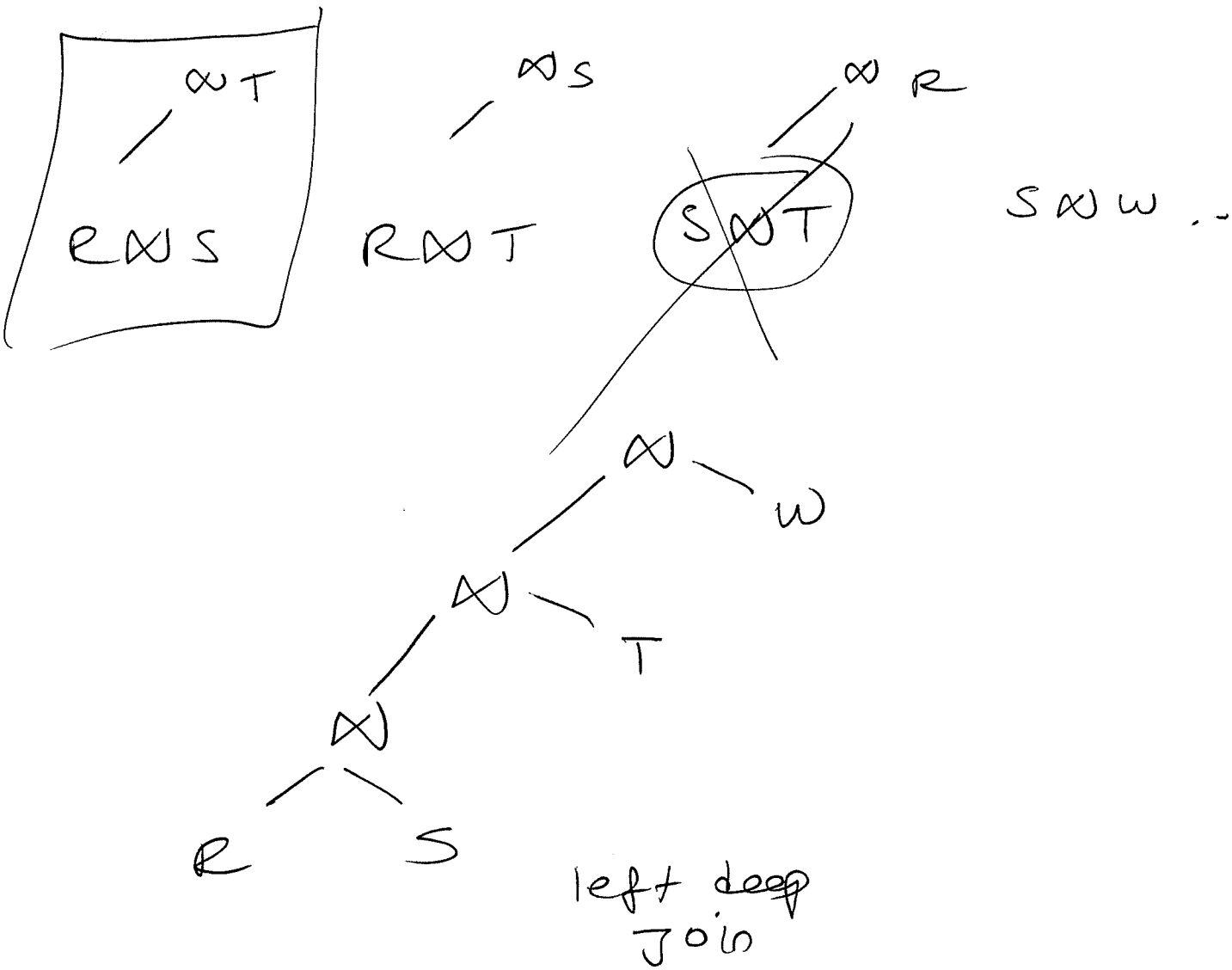
Index is 4 level

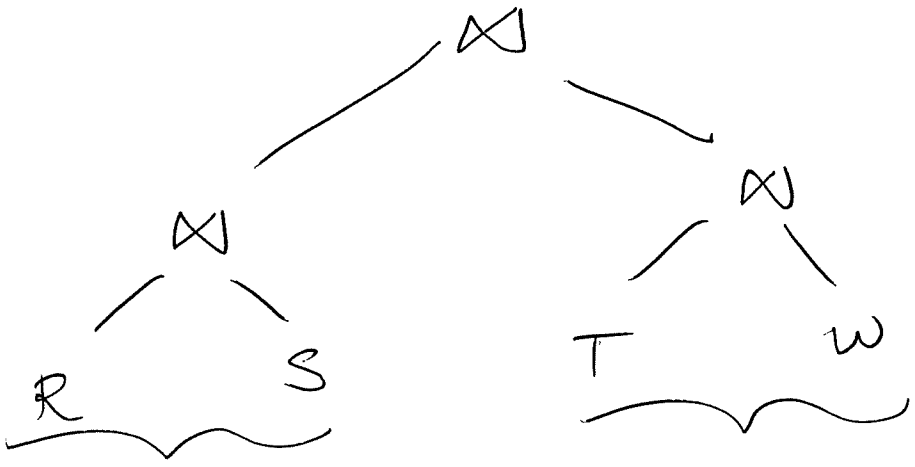
60241 leaf nodes (Scan from A/B 200/1100 — 200/4100)

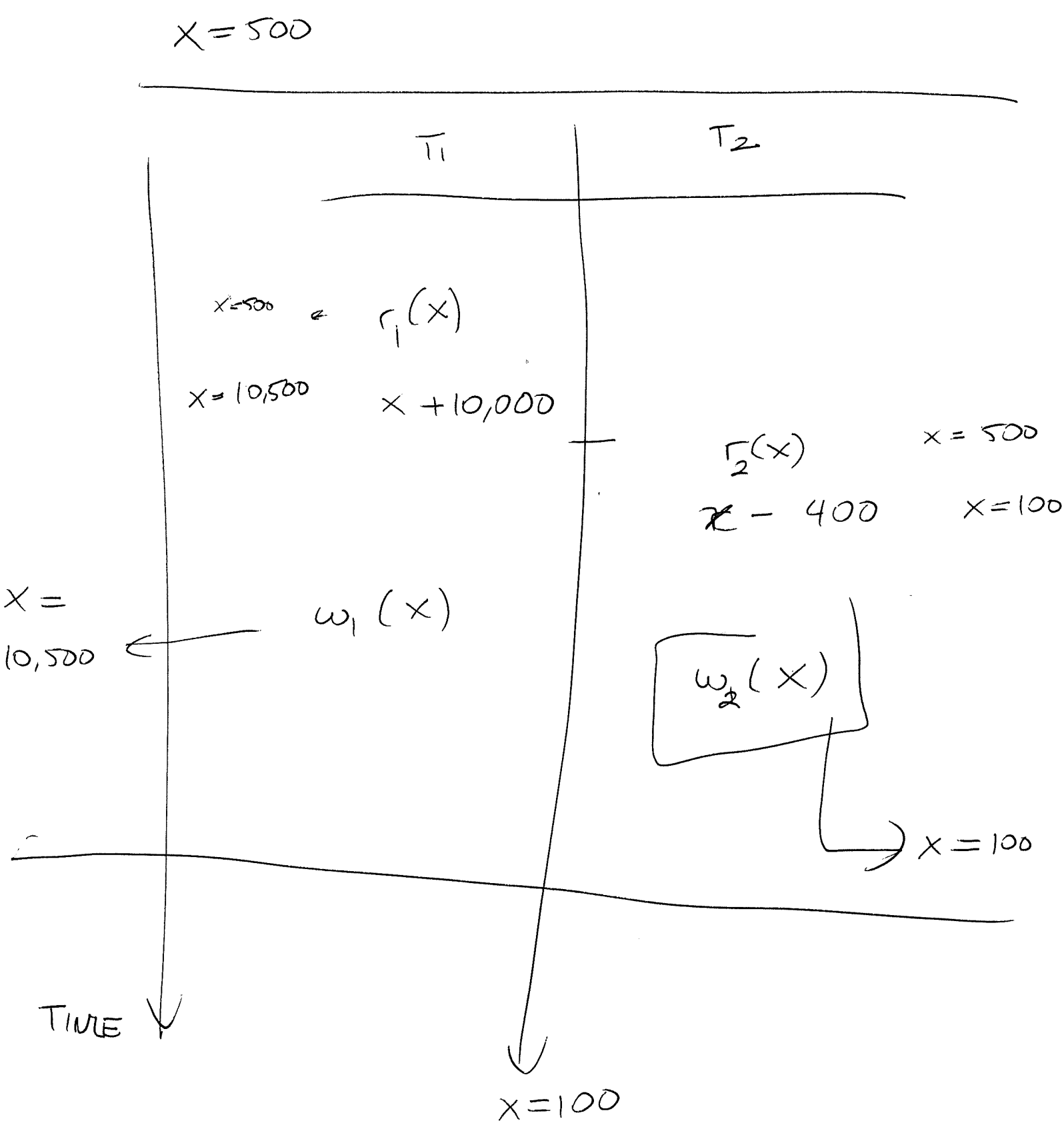
$\underbrace{3}_{\text{internal}} + \lceil \frac{15000}{166} \rceil = 94$ pages

Qopt

$R \bowtie S \bowtie T \bowtie W$




⋈
⋈
⋈
R
S
T
W


X = 500
T1
T2
X=500
r1(X)
X = 10,500
X + 10,000
r2(X)
X = 500
X - 400
X = 100
X =
10,500
ω1 (X)
ω2 (X)
X = 100
TIME
X = 100