



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



CHAI

Humanities-Centered AI

Databases

**Malte Luttermann – Institute for Humanities-Centered
Artificial Intelligence (CHAI)**

December 18, 2025

Topics for Today

- Merge Sort

Acknowledgement

- Some of the upcoming content is taken from the following lecture (and corresponding exercises) and has been translated and partially modified:
 - PD Dr. habil. Özgür L. Özçep: »Datenbanken«

Merge Sort

Assume your buffer can hold a maximum of three pages at a time and that each page in memory contains only up to two values. Further, consider the following sequence of numbers:

[39, 23, 15, 6, 86, 73, 3, 42, 91, 51, 84, 77, 25, 14, 80, 13, 72, 20, 62, 46]

Perform a merge sort (2-way merge) manually on the sequence.

Merge Sort

input file

39	23	15	6	86	73	3	42	91	51	84	77	25	14	80	13	72	20	62	46
----	----	----	---	----	----	---	----	----	----	----	----	----	----	----	----	----	----	----	----

Merge Sort

input file

39	23	15	6	86	73	3	42	91	51	84	77	25	14	80	13	72	20	62	46
----	----	----	---	----	----	---	----	----	----	----	----	----	----	----	----	----	----	----	----

1-page runs

23	39	6	15	73	86	3	42	51	91	77	84	14	25	13	80	20	72	46	62
----	----	---	----	----	----	---	----	----	----	----	----	----	----	----	----	----	----	----	----

Merge Sort

input file

39	23	15	6	86	73	3	42	91	51	84	77	25	14	80	13	72	20	62	46
----	----	----	---	----	----	---	----	----	----	----	----	----	----	----	----	----	----	----	----

1-page runs

23	39	6	15	73	86	3	42	51	91	77	84	14	25	13	80	20	72	46	62
----	----	---	----	----	----	---	----	----	----	----	----	----	----	----	----	----	----	----	----

2-page runs

6	15
23	39

3	42
73	86

51	77
84	91

13	14
25	80

20	46
62	72

Merge Sort

2-page runs

6	15
23	39

3	42
73	86

51	77
84	91

13	14
25	80

20	46
62	72

4-page runs

3	6
15	23
39	42
73	86

13	14
25	51
77	80
84	90

20	46
62	72

Merge Sort

4-page runs

3	6
15	23
39	42
73	86

13	14
25	51
77	80
84	90

20	46
62	72

8-page runs

3	6	13	14	15	23	25	39	42	51	73	77	80	84	86	91
---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----

20	46
62	72

Merge Sort

4-page runs

3	6
15	23
39	42
73	86

13	14
25	51
77	80
84	90

20	46
62	72

8-page runs

3	6	13	14	15	23	25	39	42	51	73	77	80	84	86	91
---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----

20	46
62	72

10-page runs

3	6	13	14	15	20	23	25	39	42	46	51	62	72	73	77	80	84	86	91
---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Merge Sort

How many passes were required to sort the given sequence? Please also provide the general formula for calculating the number of passes.

Merge Sort

How many passes were required to sort the given sequence? Please also provide the general formula for calculating the number of passes.

To sort n pages, the number of passes required is given by the formula

$$1 + \lceil \log_2 n \rceil.$$

In our example, 10 pages were sorted, which required

$$1 + \lceil \log_2 10 \rceil = 1 + 4 = 5$$

passes.

Merge Sort

How many passes are required to sort a sequence of 1 000 pages?

Merge Sort

How many passes are required to sort a sequence of 1 000 pages?

To sort $n = 1\,000$ pages, the number of passes required is given by

$$1 + \lceil \log_2 n \rceil = 1 + \lceil \log_2 1\,000 \rceil = 1 + 10 = 11.$$

Merge Sort

How many passes are required to sort a sequence of 1 000 000 pages?

Merge Sort

How many passes are required to sort a sequence of 1 000 000 pages?

To sort $n = 1\,000\,000$ pages, the number of passes required is given by

$$1 + \lceil \log_2 n \rceil = 1 + \lceil \log_2 1\,000\,000 \rceil = 1 + 20 = 21.$$

Merge Sort

Assume your buffer can hold a maximum of $B = 4$ pages at a time and that each page in memory contains only up to two values. Further, consider the following sequence of numbers:

[3, 4, 6, 2, 9, 4, 8, 7, 5, 9, 2, 7, 6, 1, 4, 2, 5, 6, 3, 1, 2]

Perform a merge sort (external merge sort) manually on the sequence.

Merge Sort

input file

3	4	6	2	9	4	8	7	5	9	2	7	6	1	4	2	5	6	3	1	2	
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	--

Merge Sort

input file

3	4	6	2	9	4	8	7	5	9	2	7	6	1	4	2	5	6	3	1	2	
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	--

4-page runs

2	3
4	4
6	7
8	9

1	2
2	4
5	6
7	9

1	2
3	5
6	

Merge Sort

input file

3	4	6	2	9	4	8	7	5	9	2	7	6	1	4	2	5	6	3	1	2	
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	--

4-page runs

2	3
4	4
6	7
8	9

1	2
2	4
5	6
7	9

1	2
3	5
6	

12-page runs

1	1	2	2	2	2	3	3	4	4	4	5	5	6	6	6	7	7	8	9	9	
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	--

Merge Sort (Explanation)

Initial pass:

- The file consists of $n = 12$ pages
- $B = 4$ pages are read into main memory
- The data is sorted in main memory
- Thus, n / B runs, each of size B (last run may also be smaller)

Merge pass:

- Use $B - 1$ pages of the buffer to read in pages
- Use the remaining page to produce and write the output (page by page)
- Thus, instead of $\log_2(n)$ passes required in the merge phase of two-way merge sort, we now need $\log_{B-1} n / B$ passes

Merge Sort

How many passes were required to sort the given sequence? Please also provide the general formula for calculating the number of passes.

Merge Sort

How many passes were required to sort the given sequence? Please also provide the general formula for calculating the number of passes.

To sort n pages with a buffer of B pages, the number of passes required is given by the formula

$$1 + \lceil \log_{B-1} n / B \rceil.$$

Here, $n = 12$ pages were sorted with a buffer of $B = 4$ pages, which required

$$1 + \lceil \log_{B-1} n / B \rceil = 1 + \lceil \log_3 12 / 4 \rceil = 1 + \lceil \log_3 3 \rceil = 1 + 1 = 2$$

passes.