

Exercise walk-through

File organisation and access ■ 13.2

eXercise

You must be able to

- describe **serial**, **sequential** and **random** file organisation
- match **sequential** and **direct** access to a problem
- use a **hash function** and resolve a **collision**

Method — File organisation and access

Organisation		Order of records	Best for	Access
serial		order added	logs, audit trails	sequential only
sequential		sorted by a key	master files, in-order reports	sequential / direct
random		position from the key (hash)	fast single-key lookup	direct

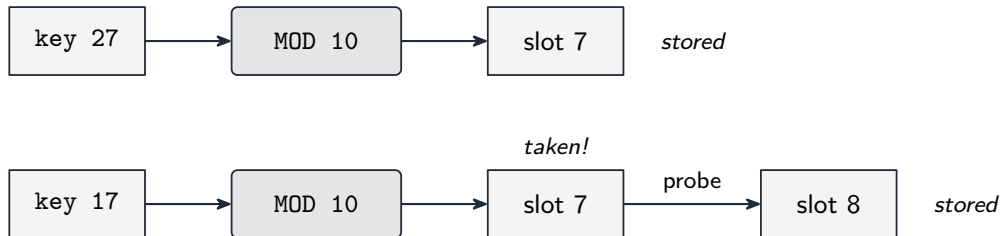
Sequential access reads from start to end; **direct access** jumps straight to a known position.

Method — Hashing

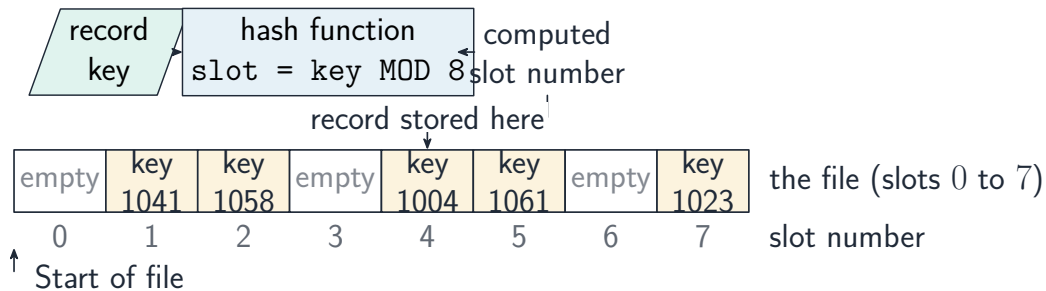
A **hash function** turns a key into an **address**: e.g. $\text{address} \leftarrow \text{key} \bmod N$.

A **collision** is when two keys hash to the same slot. Resolve it by **linear probing** (try the next free slot) or **chaining** (each slot holds a linked list). Keep the **load factor** ($\text{records} \div \text{slots}$) below $\sim 70\%$.

Method — Hashing



Method — Hashing



Random (direct) file access by address

Practice 2.1 ■ 2 marks

State the difference between a **serial** file and a **sequential** file.

Solution 2.1

Method: File organisation and access

Worked steps

A **serial** file keeps records in the **order they were added** (unsorted); a **sequential** file keeps them **sorted by a key field** **B1** **B1**.

Practice 2.2 ■ 1 mark

State which file organisation gives the **fastest single-key lookup**.

Solution 2.2

Method: File organisation and access

Worked steps

A **random (direct-access)** file [B1](#).

Practice 2.3 ■ 1 mark

State what a **hash function** does.

Solution 2.3

Method: Hashing

Worked steps

It turns a **record key** into an **address/slot** where the record is stored **B1**.

Practice 2.4 ■ 1 mark

State what a **collision** is.

Solution 2.4

Worked steps

When **two different keys** hash to the **same address/slot** **B1**.

Exam-style 3.1 ■ 3 marks

A file uses address $\leftarrow \text{key} \bmod 10$. Give the slot number for the keys 45, 23 and 30.

Solution 3.1

Method: Hashing

Worked steps

$45 \text{ MOD } 10 = 5$; $23 \text{ MOD } 10 = 3$; $30 \text{ MOD } 10 = 0$ **B1** **B1** **B1**.

Exam-style 3.2 ■ 2 marks

Keys 45 and 35 both hash to slot 5. Explain how **linear probing** stores the second one.

Solution 3.2

Method: Hashing

Worked steps

$35 \text{ MOD } 10 = 5$, but slot 5 is **taken** by 45 (B1), so linear probing tries the **next slot** (slot 6), and stores 35 there (the first free slot, wrapping if needed) (B1).

Exam-style 3.3 ■ 2 marks

State **one** advantage and **one** disadvantage of a **random (direct-access)** file.

Solution 3.3

Worked steps

Advantage: very **fast direct lookup** by key. Disadvantage: reading in **key order** is harder, and collisions must be handled (space may be wasted) **B1** **B1**.

Exam-style 3.4 ■ 2 marks

A bank must look up an account **quickly** by its account number. State the best file organisation and **justify** your choice.

Solution 3.4

Method: File organisation and access

Worked steps

A **random (direct-access)** file using a **hash** of the account number **B1**, because it gives near-instant lookup of a single record without searching the whole file **B1**.