COMP 512 - Assignment 10

## There are four major methods of organizing data files in secondary storage

### Sequential

### Indexed Sequential

### Direct

### Indexed

# Sequential File Organization

## Records are stored in order according to a key field

### If a particular record is desired, all prior records must be read first

### To update a record, a new sequential file must be created, with changed and unchanged records

## Tape storage uses sequential organization

## The Sequential File

### New records are placed in a log file or transaction file

### Batch update is performed to merge the log file with the master file

1. Indexed Sequential File Organization

## Indexed Sequential File

### New records are added to an overflow file

### Record in main file that precedes it is updated to contain a pointer to the new record

### The overflow is merged with the main file during a batch update

### Multiple indexes for the same key field can be set up to increase efficiency

1. Direct File Organization

* Also called random access
* Go directly to desired record by using a key
  + Computer does not have to read all prior records
  + Hashing algorithm used to determine address of given key
* Requires disk storage

1. Indexed File Organization

* Combines elements of sequential and direct methods
  + Records stored sequentially, but file also contains an index
  + Index stored sequentially, contains record key
  + Data accessed by record key

## Direct file organization applies the mathematical formula through hasing algorithm. Applies mathematical formula to key to determine disk address of given record. Collision occurs when hashing algorithm produces same disk address for two different keys.