# Operating Systems

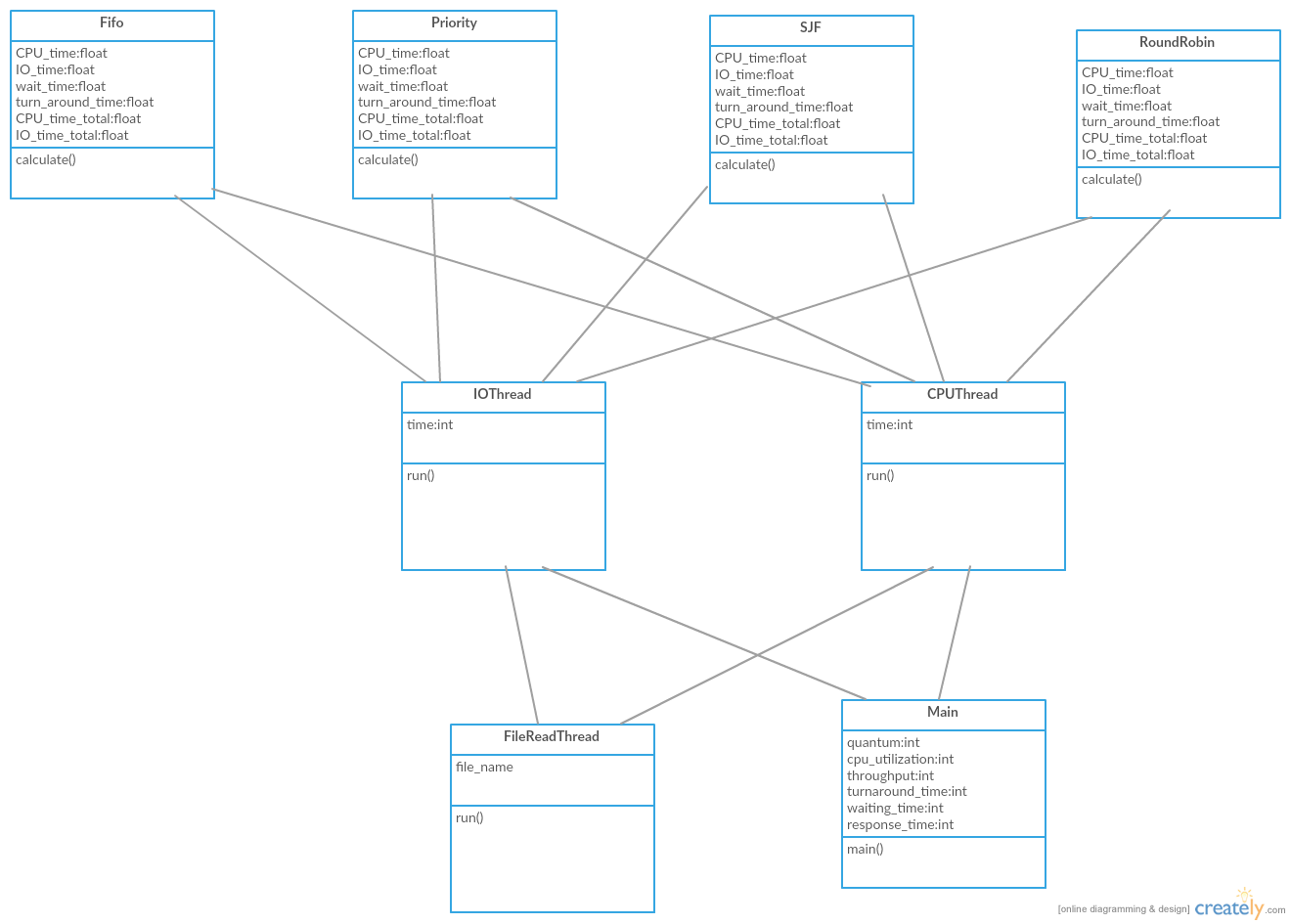
# Homework – 2

The following class diagram shows the overall structure of the program. The class IOThread and CPUThread are the CPU and Input/Output threads in the program. The classes FIFO, Priority, SJF and Roundrobin are for the following CPU scheduling algorithms.

**FIFO:** The process that requests the CPU first is allocated the CPU first. The disadvantage of this process is the average waiting time is often quite long.

**SJF:** The short job first scheduling algorithm, associates with each process the length of the process’s next CPU burst. When the CPU is available, it is assigned to the process that has the smallest next CPU burst. If the next CPU bursts of two processes are the same, FCFS scheduling is used to break the tie.

**Priority:** A priority is associated with each process, and the CPU is allocated to the process with the highest priority. Equal priority processes are scheduled in FCFS order.



**Round Robin:** The round robin scheduling algorithm is designed especially for time sharing systems. It is similar FCFS scheduling, but preemption is added to enable the system to switch in between processes. A small unit of time called quantum is defined. The ready queue is treated as a circular queue. The CPU scheduler goes around the ready queue, allocating the CPU to each process for a time interval of up to 1 time quantum.

**Performance Measures:**

**CPU utilization**: What percent of the time the CPU is used. We want to keep the CPU as busy as possible. Conceptually, CPU utilization can range from 0 to 100 percent. In a real system, it should range from 40 percent (for a lightly loaded system) to 90 percent (for a heavily loaded system).

**Throughput**: Number of processes that complete their execution per time unit. If the CPU is busy executing processes, then work is being done. One measure of work is the number of processes that are completed per time unit, called **throughput**. For long processes, this rate may be one process per hour; for short transactions, it may be ten processes per second.

**Turnaround time**: Amount of time to execute a particular process. From the point of view of a particular process, the important criterion is how long it takes to execute that process. The interval from the time of submission of a process to the time of completion is the turnaround time. Turnaround time is the sum of the periods spent waiting to get into memory, waiting in the ready queue, executing on the CPU, and doing I/O.

**Waiting time**: Amount of time a process has been waiting in the ready queue. The CPU-scheduling algorithm does not affect the amount of time during which a process executes or does I/O. It affects only the amount of time that a process spends waiting in the ready queue. Waiting time is the sum of the periods spent waiting in the ready queue.

**Response time**: Amount of time it takes from when a request was submitted until the first response is produced, not output (for time-sharing environment). In an interactive system, turnaround time may not be the best criterion. Often, a process can produce some output fairly early and can continue computing new results while previous results are beingoutput to the user. Thus, another measure is the time from the submission of a request until the first response is produced. This measure, called response time, is the time it takes to start responding, not the time it takes to output the response. The turnaround time is generally limited by the speed of the output device.