



Práctica 2 (Colas)

En un grupo de investigación necesitan gestionar el acceso a un conjunto de recursos de computación disponibles para el entrenamiento de algoritmos de IA. Los procesos que pueden querer ejecutar los investigadores se dividen entre aquellos que hacen uso intensivo de CPU o aquellos que lo hacen con GPU. Además, se pueden subdividir en dos categorías adicionales atendiendo al tiempo estimado de ejecución bajo o alto (*short / long*).

Categorías de procesos	
Tipo de recurso	Tiempo Estimado Ejecución
CPU	SHORT
	LONG
GPU	SHORT
	LONG

Para cada tipo de recurso (CPU y GPU) y tiempo estimado de ejecución de procesos se define una **cola de ejecución**: dos colas para procesos con ejecución en CPU y duración estimada baja y alta, y otras dos para procesos con ejecución en GPU y duración estimada baja y alta. Es decir, en total existen cuatro colas de ejecución con procesos en espera.

El sistema deberá simular la gestión de los procesos de las colas de ejecución, de modo que vayan ejecutándose según el orden de llegada a cada cola, teniendo en cuenta que se aplicarán **penalizaciones** a los usuarios que informen de que un proceso es de corta duración y en realidad muestre un tiempo de ejecución correspondiente con un proceso de larga duración. La penalización se aplicará en el siguiente proceso del usuario que llegue a cualquiera de las **colas de ejecución**.

Además, existirá una **única cola de registro de entrada de procesos** en la que inicialmente se añaden todos los procesos de los usuarios en orden de llegada. A partir de esta **cola de registro**, el sistema irá distribuyendo los procesos en las **colas de ejecución** correspondientes (actualmente las 4 descritas previamente).

Tareas

Se simulará la tarea de gestión y ejecución de procesos ayudándonos de varias estructuras de datos (al menos las siguientes):

- Una clase que represente a un **proceso** con todos los atributos necesarios (identificador del proceso, identificador del usuario, tipo de recurso (cpu / gpu), tiempo estimado de ejecución (long / short), tiempo de inicio de ejecución, etc. Se incluirán todos los atributos que se consideren necesarios.

- Una cola para el registro de entrada de procesos de los usuarios (**cola de registro**).
- Una clase **gestor de colas**, para la gestión de las cuatro colas de ejecución, para la gestión de las penalizaciones aplicadas a los usuarios y para el almacenamiento temporal de los procesos que en un momento dado están ejecutándose en cada recurso (cpu-short, cpu-long, gpu-short, gpu-long).

El proceso de simulación que debe implementarse consta de los pasos siguientes:

1. Leer de un fichero los datos de los procesos de los usuarios, que se utilizarán para simulación, e insertarlos en una cola de registro de entrada de procesos. El fichero contiene la información de cada proceso en una línea, con 5 campos separados por espacios:
 - i. Identificador único del proceso.
 - ii. Identificador único del usuario propietario del proceso.
 - iii. Tipo de proceso. Los valores posibles son: cpu / gpu.
 - iv. Tiempo estimado de ejecución indicado por el usuario. Los valores posibles son: short y long.
 - v. Tiempo de ejecución. Valor entero que indica las unidades de tiempo que requerirá el proceso para ejecutarse en la simulación. Este valor se incluye únicamente para simplificar la simulación y facilitar la verificación de la correcta ejecución de esta.
2. Inicializar un contador (numérico) que simulará el tiempo transcurrido durante la simulación.
3. Repetir hasta que no queden elementos en la *cola de registro* de procesos, no haya elementos en ninguna de las *colas de ejecución*, y todos los procesos hayan terminado su ejecución:
 - a. Incrementar el contador de tiempo en una unidad.
 - b. Si la *cola de registro* no está vacía, extraer el primer proceso, asignarle el tiempo actual como *tiempo de entrada* y añadirlo a la clase *gestor de colas*, de modo que se inserte en la cola adecuada según su *tipo (cpu/gpu)* y *tiempo estimado (short/long)*.
 - c. Si el *gestor de colas* tiene almacenado algún proceso que ya está siendo ejecutado, comprobar si se ha cumplido su tiempo de ejecución ($t_{actual} \geq t_{inicio_ejecución} + t_{ejecución}$) y en ese caso darlo por terminado y eliminarlo.
 - d. Cuando un proceso termina su ejecución, si había sido definido por el usuario como un proceso de corta duración (*short*) y su tiempo de ejecución es mayor que **5 unidades de tiempo**, el *gestor de ejecución* registrará que ese usuario tiene una **penalización activa**.
 - e. Procesar todas las *colas de ejecución* (2 colas para CPU y 2 colas para GPU) de modo que:
 - i. Se extrae el proceso al frente de cada una de las colas de ejecución y se inicia su ejecución. Esto se simula estableciendo el tiempo actual como *tiempo de inicio de ejecución* de ese proceso y haciendo que el *gestor de colas* guarde dicho proceso hasta que finalice su ejecución.

- ¡IMPORTANTE! Para cada cola, esta acción solo se realizará si el *gestor de colas* NO tiene ya registrado como proceso en ejecución un proceso previo del mismo tipo.
- ii. Si el proceso al frente de una cola *short* pertenece a un *usuario con penalización activa*, no se iniciará su ejecución sino que se eliminará de su cola y se añadirá al final de la cola (tipo cpu o gpu según corresponda) de mayor duración, *long* en este caso. Se desactivará la penalización del usuario.

El programa mostrará por pantalla en cada iteración el contador de tiempo y un mensaje en las situaciones:

- Se añade proceso a cola de ejecución, de acuerdo al formato siguiente:

Proceso añadido a cola de ejecución: <TActual><IDProceso><IDUsuario>
<Tipo><TiempoEstimadoEjecución>

- Fin de la ejecución de un proceso, de acuerdo al formato siguiente:

Proceso terminado: <TActual>><IDProceso><IDUsuario><Tipo>
<TiempoEstimadoEjecución><TEntrada><TInicioEjecución><TEjecución>

- Se registra penalización activa para un usuario, de acuerdo al formato siguiente:

Penalización activa: <TActual>><IDUsuario>

- Se aplica penalización a un proceso y se desactiva penalización, de acuerdo al formato siguiente:

Penalización aplicada: <TActual><IDProceso><IDUsuario>

Al final de la simulación presentar las siguientes estadísticas (se debe utilizar la librería [pandas](#) para el análisis de los datos):

1. Una tabla con el número medio de procesos a los que se ha aplicado una penalización, agrupado por usuario.
2. Una tabla que muestre el tiempo medio de permanencia de los procesos en el *gestor de colas*, agrupado por cada cola de ejecución (entendido como la diferencia entre el tiempo en que se inicia la ejecución de un proceso y el tiempo de entrada a la cola de ejecución).

Las estructuras de datos deberán implementarse respetando los principios básicos de orientación a objetos (herencia, encapsulación y polimorfismo). Para la estructura tipo cola debe utilizarse la implementación vista en los contenidos teóricos de la asignatura.

Entrega

Se entregará un **archivo zip** con todo el código fuente (comentado con # y clases/métodos con docstrings), y un **documento PDF** que incluya un breve manual de usuario y una breve pero exhaustiva descripción de las fases de desarrollo realizadas, que incluya las clases implementadas, y los resultados obtenidos tras la simulación.

IMPORTANTE:

- En cada archivo y en la primera página del PDF se indicará el nombre y login de ambos miembros de la pareja de prácticas.
- Los nombres de carpetas y ficheros de código .py seguirán la convención de nombres *snake-case* (palabras unidas por barra baja, sin espacios y en minúsculas, ej. `mi_fichero.py`).

Fecha límite de entrega: **miércoles 27 de marzo de 2024 a las 23:55.**

Dónde se entrega: en el apartado **Práctica 2** del Campus Virtual.

Quién entrega: sólo uno de los miembros de la pareja deberá entregar la práctica.