Administrivia

- Reminder: Homework 3 due Wednesday.
- Review sheet for midterm on the Web. Review in class Friday.

Slide 1

Dining Philosophers Problem — Review

- Scenario:
 - Five philosophers sitting around a table, each alternating between thinking and eating.
 - Between every pair of philosophers, a fork; philosopher must have two forks to eat.
 - So, neighbors can't eat at the same time, but non-neighbors can.
- Why is this interesting or important? It's a simple example of something more complex than mutual exclusion — multiple shared resources (forks), processes (philosophers) must obtain two resources together. (Why five? smallest number that's "interesting".)

Dining Philosophers — Dijkstra Solution

• Solution uses shared variables to track state of philosophers and semaphores to synchronize:

- Array of five state variables (states [5]), possible values thinking, hungry, eating. Initially all thinking.
- Semaphore mutex, initial value 1, to enforce one-at-a-time access to states.
- Array of five semaphores self[5], initial values 0, to allow us to make philosophers wait.
- And then the code is somewhat complex ...

Dining Philosophers — Code

• Shared variables as on previous slide.

Pseudocode for philosopher i:

```
while (true) {
    think();
    down(mutex);
    state[i] = hungry;
    test(i);
    up(mutex);
    down(self[i]);
    eat();
    down(mutex);
    state[i] = thinking;
    test(right(i));
    test(left(i));
    up(mutex);
```

Pseudocode for function:

Slide 4

Dining Philosophers — Dijkstra Solution Works?

- Could there be problems with access to shared state variables?
- Do we guarantee that neighbors don't eat at the same time?
- Do we allow non-neighbors to eat at the same time?
- Could we deadlock?
 - Does a hungry philosopher always get to eat eventually?

Dining Philosophers — Dijkstra Solution Works?

- Could there be problems with access to shared state variables? No (because all accesses are "protected" by mutual-exclusion semaphore).
- Do we guarantee that neighbors don't eat at the same time? Yes.
- Do we allow non-neighbors to eat at the same time? Yes.
- Could we deadlock? No.
- Does a hungry philosopher always get to eat eventually? Usually. Exception is when two next-to-neighbors (e.g., 1 and 3) seem to conspire to starve their common neighbor (e.g., 2).

Slide 5

Dining Philosophers — Chandy/Misra Solution

 Original solution allows for scenarios in which one philosopher "starves" because its neighbors alternate eating while it remains hungry.

Briefly, we could improve this by maintaining a notion of "priority" between
neighbors, and only allow a philosopher to eat if (1) neither neighbor is eating,
and (2) it doesn't have a higher-priority neighbor that's hungry. After a
philosopher eats, it lowers its priority relative to its neighbors.

Slide 7

Other Classical Problems

- Readers/writers (in textbook).
- $\bullet\,$ Sleeping barber, drinking philosophers, \ldots
- Advice if you ever have to solve problems like this "for real", read the literature . . .

Review — Processes and Context Switches

• Recall idea behind process abstraction — make every activity we want to manage a "process", and run them "concurrently".

- Apparent concurrency provided by interleaving. (Some) true concurrency provided by multiple cores/processors.
- To make this work process table, ready/running/blocked states, context switches.
- Context switches triggered by interrupts I/O, timer, system call, etc.
- On interrupts, interrupt handler processes interrupt, and then goes back to some process but which one?

Which Process To Run Next?

- Deciding what process to run next scheduler/dispatcher, using "scheduling algorithm".
- When to make scheduling decisions?
 - When a new process is created.
 - When a running process exits.
 - When a process becomes blocked (I/O, semaphore, etc.).
 - After an interrupt.
- One possible decision "go back to interrupted process" (e.g., after I/O interrupt).

Slide 9

Scheduler Goals

- Importance of scheduler can vary; extremes are
 - Single-user system often only one runnable process, complicated decision-making may not be necessary (though still might sometimes be a good idea).
 - Mainframe system many runnable processes, queue of "batch" jobs waiting, "who's next?" an important question.
 - Servers / workstations somewhere in the middle.
- First step is to be clear on goals want to make "good decisions", but what does that mean? Typical goals for any system:
 - Fairness similar processes get similar service.
 - Policy enforcement "important" processes get better service.
 - Balance all parts of system (CPU, I/O devices) kept busy (assuming there is work for them).

Aside — Terminology

- Discussion often in term of "jobs" holdover from mainframe days, means "schedulable piece of work".
- Processes usually alternate between "CPU bursts" and I/O, can be categorized as "compute-bound" ("CPU-bound") or "I/O bound".

• Scheduling can be "preemptive" or "non-preemptive".

Slide 12

Scheduler Goals By System Type

- For batch (non-interactive) systems, possible goals (might conflict):
 - Maximize throughput jobs per hour.
 - Minimize turnaround time.
 - Maximize CPU utilization.

Preemptive scheduling may not be needed.

- For interactive systems, possible goals:
 - Minimize response time.
 - Make response time proportional (to user's perception of task difficulty).

Preemptive scheduling probably needed.

- For real-time systems, possible goals:
 - Meet time constraints/deadlines.
 - Behave predictably.

Scheduling Algorithms

- Many, many scheduling algorithms, ranging from simple to not-so-simple.
- Point of reviewing lots of them? notice how many ways there are to solve the same problem ("who should be next?"), strengths/weaknesses of each.

Slide 14

Minute Essay

• What was interesting about Homework 2? what was interesting?