CS370 Operating Systems
Colorado State University
Yashwant K Malaiya
Fall 2022 L12
Synchronization (Chap 6, 7)

Slides based on
• Text by Silberschatz, Galvin, Gagne
• Various sources
Notes: we are using the terms in a generic way. There are specific implementations for POSIX and Java.

• **Atomic instructions**: Hardware (Assembly/Machine), not *high level* like C

• **Mutex (0 or 1)**: for mutual exclusion (lock). Owned by the locking process which acquires/releases by
  – `wait( )` get the resource or join the waiting list
  – `signal( )` release the resource, and wake up a process

• **Semaphores (any integer value)**: general, may be used for counting resources/waiting processes. Shared. Applicable to different types of synchronization problems.
  – 0: no waiting threads
  – Positive: no waiting threads, a wait operation would not put the invoking thread in queue.
  – Negative: number of processes/threads waiting

• **Semaphore implementation**
  – Hardware/software implementations to ensure `wait()` and `signal( )` atomic.

• **Semaphore usage**: see POSIX/Java documentation. [Example](#).
Project

- A. Research
- B. Development

Deliverable *D1 Team composition and idea proposal* specified separately in the document. Similarly, D2, D3, D4 and D5 are specified.

You have to do some research for both of them.
Specific sources: database indexes

- Google Scholar
  - Forward links: Paper X Cited by
  - Backward Links: Paper X cites

- Researcher sites
  - Personal/Group Website
  - DBLP
  - Google Scholar: researcher

- CSU Library etc.

General (accessible through CSU Library)

- ACM Digital Library
- IEEEXplore Digital Library
- ScienceDirect etc
Research: Source types

• Journals: published several times a year
  – Rigorously reviewed, long publication delay
  – Journal, Transactions, ...

• Conferences: held once a year, proceedings published
  – Conference, Symposium, ...

• Research groups
  – Industry, academic, consultants: web site

• News, Industry publications
  – Magazines, blogs, white papers, product website

• Books: often well-known stuff
Research: How to Read a Paper: THE THREE-PASS APPROACH

• The first pass: Read
  – the title, abstract, and introduction
  – section and sub-section headings, but ignore everything else
  – the conclusions

• The second pass: Read
  – figures, diagrams and other illustrations
  – mark relevant unread references for further reading
  – Do you need to read it in detail?

• The third pass: Read critically
  – identify and challenge assumption and views
  – Loop up references needed

Research: Avoid Prior Bias

Look, half the work is done! All you need to do is fill in the top part so we can legally say the bottom part.

DATA:

CONCLUSION: Eating chocolate will make you look younger and thinner.

© Wiley Ink, inc./Distributed by Universal Uclick via Cartoonstock
Evaluating research

• These are the attributes generally evaluated
• Novelty/interest/Applicability
• Technical: Extent of research/contribution
  – Key sources? Recent developments?
  – citations
• Presentation
  – Visuals: Non-text: diagrams, charts, algorithms
  – Systematic/quantitative: tables, numbers
  – Readability, coherence
• Overall
Classical problems

– Bounded-Buffer Problem
– Readers and Writers Problem
– Dining-Philosophers Problem

• Bounded buffer Review
  – n buffers, each can hold one item
  – A binary semaphore: mutex
    • Provides mutual exclusion for accesses to buffer pool
    • Initialized to 1
  – Two counting semaphores
    • empty: Number of empty slots available, Initialized to n
    • full: Number of filled slots available n, Initialized to 0
Readers-Writers Problem

• A data set is shared among a number of concurrent processes
  – Readers – only read the data set; they do not perform any updates
  – Writers – can both read and write

• Problem
  – allow multiple readers to read at the same time
  – Only one single writer can access the shared data at the same time. No readers permitted when writer is accessing the data.

• Several variations of how readers and writers are considered – all involve some form of priorities
Readers-Writers Problem

- Shared Data
  - Data set
  - Semaphore `rw_mutex` initialized to 1 (mutual exclusion for writer)
  - Semaphore `mutex` initialized to 1 (mutual exclusion for `read_count`)
  - Integer `read_count` initialized to 0 (how many readers?)
• The structure of a writer process

```c
do {
    wait(rw_mutex);
    ...
    /* writing is performed */
    ...
    signal(rw_mutex);
} while (true);
```
The structure of a reader process

```c
do {
    wait(mutex);
    read_count++;
    if (read_count == 1)
        wait(rw_mutex);
    signal(mutex);
    /* reading is performed */
    ...
    wait(mutex);
    read_count--;
    if (read_count == 0)
        signal(rw_mutex);
    signal(mutex);
} while (true);
```

First reader needs to wait for the writer to finish. If other readers are already reading, a new reader process just goes in.

Mutex for mutual exclusion to read_count

When: writer in critical section and if n readers waiting 1 is queued on rw_mutex (n-1) queued on mutex

Cannot read if writer is writing

When the last reader leaves, a writer can go in.
Readers-Writers Problem Variations

• **First** variation – no reader kept waiting unless writer has already obtained permission to use shared object

• **Second** variation – once writer is ready, it performs the write ASAP, i.e. if a writer is waiting, no new readers may start.

• Both may have starvation leading to even more variations

• Problem is solved on some systems by kernel providing reader-writer locks
Dining-Philosophers Problem

- Philosophers spend their lives alternating thinking and eating
- Don’t interact with their neighbors, occasionally try to pick up 2 chopsticks (one at a time) to eat from bowl
  - Need both to eat,
  - then release both when done
- Each chopstick is a semaphore
  - Grab by executing `wait()`
  - Release by executing `signal()`
- Shared data
  - Bowl of rice (data set)
  - Semaphore `chopstick[5]` initialized to 1
Dining-Philosophers Problem

Plato, Confucius, Socrates, Voltaire and Descartes
Dining-Philosophers Problem Algorithm: Simple solution?

• The structure of Philosopher $i$:
  
  ```c
  do {
    wait (chopstick[i] )
    wait (chopStick[(i + 1) % 5] )
    // eat
    signal (chopstick[i] )
    signal (chopstick[(i + 1) % 5] )
    // think
  } while (TRUE);
  ```

• What is the problem with this algorithm?
  – If all of them pick up the the left chopstick first - Deadlock
• Deadlock handling
  – Allow at most 4 philosophers to be sitting simultaneously at the table (with the same 5 forks).
  – Allow a philosopher to pick up the forks only if both are available (picking must be done in a critical section.
  – Use an asymmetric solution -- an odd-numbered philosopher picks up first the left chopstick and then the right chopstick. Even-numbered philosopher picks up first the right chopstick and then the left chopstick.
Problems with Semaphores

• Incorrect use of semaphore operations:
  – Omitting of wait (mutex)
    • Violation of mutual exclusion
  – or signal (mutex)
    • Deadlock!

• Solution:
  – Monitors: a higher-level implementation of synchronization
Related classes

Classes that follow CS370

- CS455 Distributed Systems  Spring
- CS457 Networks  Fall
- CS470 Computer Architecture  Spring
- CS475 Parallel Programming  Fall
- CS435: Introduction to Big Data  Spring
Monitors

**Monitor**: A high-level abstraction that provides a convenient and effective mechanism for process synchronization

- Abstract data type, internal variables only accessible by code within the procedure
- Only one process may be active within the monitor at a time
  - Automatically provide mutual exclusion
  - Implement waiting for conditions
- Queues:
  - for entry
  - for each condition
- Originally proposed for Concurrent Pascal 1975
- Directly supported by Java (see self exercise) but not C
• Only one process may be actively under execution in the monitor.
• A generic monitor construct is used here. Implementation varies by language.

```plaintext
monitor monitor-name
{
    // shared variable declarations

    procedure P1 (...) { .... }

    procedure Pn (...) {......}

    Initialization code (...) { ... }
}
}
Preliminary Schematic view of a Monitor

- Only one process/thread in the Monitor
  - Provides an easy way to achieve mutual exclusion

But ... we also need a way for processes to block when they cannot proceed.
  - Refinement next ...

Shows 4 processes waiting in the queue.
Condition Variables

Some actions need some conditions to go ahead.
The **condition** construct

• `condition x, y;`

• Two operations are allowed on a condition variable:
  - `x.wait()` — a process that invokes the operation is suspended until `x.signal()`
  - `x.signal()` — resumes one of processes (if any) that invoked `x.wait()`
    • If no `x.wait()` on the condition variable, then it has no effect on the variable. *Signal is lost.*
Difference between the signal() in semaphores and monitors

• Condition variables in Monitors: Not persistent
  – If a signal is performed and no waiting threads?
    • Signal is simply ignored
  – During subsequent wait operations
    • Thread (or process) blocks

• Compare with semaphores
  – Signal increments semaphore value even if there are no waiting threads
    • Future wait operations would immediately succeed!
Monitor with Condition Variables

- Entry queue
- Shared data
- Queues associated with $x, y$ conditions
- Operations
- Initialization code
Condition Variables Choices

• If process P invokes `x.signal()`, and process Q is suspended in `x.wait()`, what should happen next?
  – Both Q and P cannot execute in parallel. If Q is resumed, then P must wait

• Options include
  – Signal and wait – P waits until Q either leaves the monitor or it waits for another condition
  – Signal and continue – Q waits until P either leaves the monitor or it waits for another condition
  – Both have pros and cons – language implementer can decide
  – Monitors implemented in Concurrent Pascal (‘75) compromise
    • P executing signal immediately leaves the monitor, Q is resumed
    • Implemented in other languages including C#, Java
enum {THINKING, HUNGRY, EATING} state[5];

• state[i] = EATING only if
  – state[(i+4)%5] != EATING && state[(i+1)%5] != EATING

• condition self[5]
  – Delay self when HUNGRY but unable to get chopsticks

Sequence of actions

• Before eating, must invoke pickup()
  – May result in suspension of philosopher process
  – After completion of operation, philosopher may eat

    think
    DiningPhilosophers.pickup(i);
    eat
    DiningPhilosophers.putdown(i);
    think
enum {THINKING, HUNGRY, EATING} state[5];

Can I eat? If not, I’ll wait
The pickup() and putdown() operations

```c
// Definition of the DiningPhilosophers monitor

monitor DiningPhilosophers
{
    enum { THINKING, HUNGRY, EATING} state [5];
    condition self [5];

    void pickup (int i) {
        state[i] = HUNGRY;
        test(i);  //below
        if (state[i] != EATING) self[i].wait;
    }

    void putdown (int i) {
        state[i] = THINKING;
        // test left and right neighbors
        test(((i + 4) % 5);
        test(((i + 1) % 5);
    }

    void test (int i) {
        if ((state[(i + 4) % 5] != EATING) &&
            (state[i] == HUNGRY) &&
            (state[(i + 1) % 5] != EATING) ) {
            state[i] = EATING ;
            self[i].signal () ;
        }
    }

    initialization_code() {
        for (int i = 0; i < 5; i++)
            state[i] = THINKING;
    }
}
```

- **Suspend self if unable to acquire chopstick**
- **Eat only if HUNGRY and Person on Left AND Right are not eating**
- **Check to see if person on left or right can use the chopstick**
- **Signal a process that was suspended while trying to eat**
Possibility of starvation

- Philosopher i can starve if eating periods of philosophers on left and right overlap
- Possible solution
  - Introduce new state: STARVING
  - Chopsticks can be picked up if no neighbor is starving
    - Effectively wait for neighbor’s neighbor to stop eating
    - REDUCES concurrency!
For each monitor
• Semaphore mutex initialized to 1
• Process must execute
  – wait(mutex) : Before entering the monitor
  – signal(mutex): Before leaving the monitor
• If several processes queued on condition x, and x.signal() is executed, which should be resumed?

• FCFS frequently not adequate

• `conditional-wait` construct of the form `x.wait(c)`
  – Where c is `priority number`
  – Process with lowest number (highest priority) is scheduled next
Allocate a single resource among competing processes using priority numbers that specify the maximum time a process plans to use the resource:

```java
R.acquire(t);
...
access the resource;
...
R.release;
```

Where R is an instance of type `ResourceAllocator`

A monitor based solution next.
A Monitor to Allocate Single Resource

```java
monitor ResourceAllocator {
    boolean busy;
    condition x;

    void acquire(int time) {
        if (busy)
            x.wait(time);
        busy = TRUE;
    }

    void release() {
        busy = FALSE;
        x.signal();
    }

    initialization code() {
        busy = FALSE;
    }
}
```

Sleep, Time used to prioritize waiting processes

Wakes up one of the processes
For simple synchronization, Java provides the `synchronized` keyword

- Synchronizing methods
  ```java
  public synchronized void increment() {
      c++;
  }
  ```
- Synchronizing blocks
  ```java
  synchronized(this) {
      lastName = name;
      nameCount++;
  }
  ```

- `wait()` and `notify()` allows a thread to wait for an event. A call to `notifyAll()` allows all threads that are on `wait()` with the same lock to be notified.

- `notify()` notifies one thread from a pool of identical threads, `notifyAll()` when threads have different purposes.

- For more sophisticated locking mechanisms, starting from Java 5, the package `java.concurrent.locks` provides additional capabilities.
Java Synchronization

Each object automatically has a monitor (mutex) associated with it

- When a method is synchronized, the runtime must obtain the lock on the object's monitor before execution of that method begins (and must release the lock before control returns to the calling code)

wait() and notify() allows a thread to wait for an event.

- **wait():** Causes the current thread to wait until another thread invokes the notify() method or the notifyAll() method for this object.
- **notify():** Wakes up a single thread that is waiting on this object's monitor. If any threads are waiting on this object, one of them is chosen to be awakened.
- A call to notifyAll() allows all threads that are on wait() with the same lock to be released, they will run in sequence according to priority.

https://www.baeldung.com/java-wait-notify
<table>
<thead>
<tr>
<th>Java Synchronization: Dining Philosophers</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>public synchronized void pickup(int i) throws InterruptedException {</td>
<td>private synchronized void test(int i) {</td>
</tr>
<tr>
<td>setState(i, State.HUNGRY);</td>
<td>if (state[left(i)] != State.EATING &amp;&amp;</td>
</tr>
<tr>
<td>test(i);</td>
<td>state[right(i)] != State.EATING &amp;&amp;</td>
</tr>
<tr>
<td>while (state[i] != State.EATING) {</td>
<td>state[i] == State.HUNGRY)</td>
</tr>
<tr>
<td>this.wait();</td>
<td>{</td>
</tr>
<tr>
<td>// Recheck condition in loop,</td>
<td>setState(i, State.EATING);</td>
</tr>
<tr>
<td>// since we might have been notified</td>
<td>// Wake up all waiting threads</td>
</tr>
<tr>
<td>// when we were still hungry</td>
<td>this.notifyAll();</td>
</tr>
<tr>
<td>}</td>
<td>}</td>
</tr>
<tr>
<td>}</td>
<td>}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>public synchronized void putdown(int i) {</td>
<td></td>
</tr>
<tr>
<td>setState(i, State.THINKING);</td>
<td></td>
</tr>
<tr>
<td>test(right(i));</td>
<td></td>
</tr>
<tr>
<td>test(left(i));</td>
<td></td>
</tr>
<tr>
<td>}</td>
<td></td>
</tr>
</tbody>
</table>
Synchronization Examples

- Solaris
- Windows
- Linux
- Pthreads
Solaris Synchronization

- Implements a variety of locks to support multitasking, multithreading (including real-time threads), and multiprocessing
- Uses **adaptive mutexes** for efficiency when protecting data from short code segments
  - Starts as a standard semaphore spin-lock
  - If lock held, and by a thread running on another CPU, spins
  - If lock held by non-run-state thread, block and sleep waiting for signal of lock being released
- Uses **condition variables**
- Uses **readers-writers** locks when longer sections of code need access to data
- Uses **turnstiles** to order the list of threads waiting to acquire either an adaptive mutex or reader-writer lock
  - Turnstiles are per-lock-holding-thread, not per-object
- Priority-inheritance per-turnstile gives the running thread the highest of the priorities of the threads in its turnstile
Windows Synchronization

• Uses interrupt masks to protect access to global resources on uniprocessor systems
• Uses spinlocks on multiprocessor systems
  – Spinlocking-thread will never be preempted
• Also provides dispatcher objects user-land which may act mutexes, semaphores, events, and timers
  – Events
    • An event acts much like a condition variable
  – Timers notify one or more thread when time expired
  – Dispatcher objects either signaled-state (object available) or non-signaled state (thread will block)
Linux Synchronization

• Linux:
  – Prior to kernel Version 2.6, disables interrupts to implement short critical sections
  – Version 2.6 and later, fully preemptive

• Linux provides:
  – Semaphores
  – atomic operations on integers
  – spinlocks
  – reader-writer versions of both

• On single-cpu system, spinlocks replaced by enabling and disabling kernel preemption
Pthreads Synchronization

- Pthreads API is OS-independent
- It provides:
  - mutex locks
  - condition variable
- Non-portable extensions include:
  - read-write locks
  - spinlocks
Alternative Approaches

• Transactional Memory
• OpenMP
• Functional Programming Languages
• A **memory transaction** is a sequence of read-write operations to memory that are performed atomically without the use of locks.

```c
void update()
{
    atomic{
        /* modify shared data*/
    }
}
```

May be implemented by hardware or software.
OpenMP

- OpenMP is a set of compiler directives and API that support parallel programming.

```c
void update(int value)
{
    #pragma omp critical
    {
        count += value
    }
}
```

The code contained within the `#pragma omp critical` directive is treated as a critical section and performed atomically.
Deadlock

Slides based on
- Text by Silberschatz, Galvin, Gagne
- Various sources
Chapter 8: Deadlocks

• System Model
• Deadlock Characterization
• Methods for Handling Deadlocks
  – Deadlock Prevention
  – Deadlock Avoidance resource-allocation
  – Deadlock Detection
  – Recovery from Deadlock
• Can you give a real life example of a deadlock?
A Kansas Law

• Early 20th century Kansas Law
  – “When two trains approach each other at a crossing, both shall come to a full stop and neither shall start up again until the other has gone”

• Story of the two silly goats: Aesop 6th cent BCE?
A contemporary example
System Model

- System consists of resources
- Resource types $R_1, R_2, \ldots, R_m$
  
  *CPU cycles, memory space, I/O devices*
- Each resource type $R_i$ has $W_i$ instances.
- Each process utilizes a resource as follows:
  - request
  - use
  - release
Resource-Allocation Graph

A set of vertices $V$ and a set of edges $E$.

- $V$ is partitioned into two types:
  - $P = \{P_1, P_2, \ldots, P_n\}$, the set consisting of all the processes in the system
  - $R = \{R_1, R_2, \ldots, R_m\}$, the set consisting of all resource types in the system

- **request edge** – directed edge $P_i \rightarrow R_j$
- **assignment edge** – directed edge $R_j \rightarrow P_i$
• Process

• Resource Type with 4 instances

• $P_i$ requests instance of $R_j$

• $P_i$ is holding an instance of $R_j$
Example of a Resource Allocation Graph

If the graph contains no cycles, then no process in the system is deadlocked. If the graph does contain a cycle, then a deadlock may exist.

Does a deadlock exist here?

P3 will eventually be done with R3, letting P2 use it.

Thus P2 will be eventually done, releasing R1. ...

P1 holds an instance of R2, and is requesting R1 ..