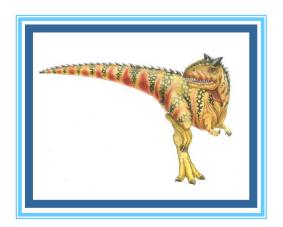
Chapter 7: Deadlocks

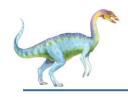




Chapter 7: Deadlocks

- System Model
- Deadlock Characterization
- Methods for Handling Deadlocks
- Deadlock Prevention
- Deadlock Avoidance
- Deadlock Detection
- Recovery from Deadlock



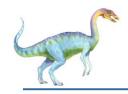


Chapter Objectives

To develop a description of deadlocks, which prevent sets of concurrent processes from completing their tasks

To present a number of different methods for preventing or avoiding deadlocks in a computer system





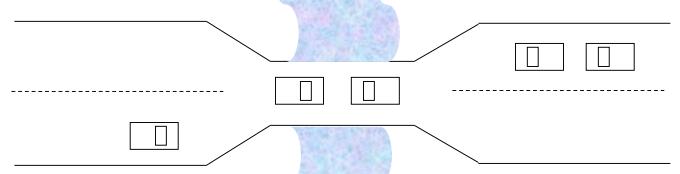
The Deadlock Problem (沈庭宇)

- A set of blocked processes each holding a resource and waiting to acquire a resource held by another process in the set
- Example
 - System has 2 disk drives
 - P₁ and P₂ each hold one disk drive and each needs another one
- Example
 - semaphores A and B, initialized to 1

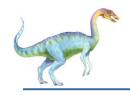
$$P_0$$
 P_1 wait (A); wait(B) wait (B); wait (A)







- Traffic only in one direction
- Each section of a bridge can be viewed as a resource
- If a deadlock occurs, it can be resolved if one car backs up (preempt resources and rollback)
- Several cars may have to be backed up if a deadlock occurs
- Starvation is possible
- Note Most OSes do not prevent or deal with deadlocks



System Model(劉秋志)

- System consists of resources
- Resource types $R_1, R_2, ..., R_m$ 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



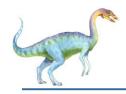


Deadlock Characterization (1/2)

Deadlock can arise if four conditions hold simultaneously.

- 1. Mutual exclusion: only one process at a time can use a resource
- 2. Hold and wait: a process holding at least one resource is waiting to acquire additional resources held by other processes
- 3. No preemption: a resource can be released only voluntarily by the process holding it, after that process has completed its task





Deadlock Characterization (2/2)

Deadlock can arise if four conditions hold simultaneously.

- **4. Circular wait:** there exists a set $\{P_0, P_1, ..., P_n\}$ of waiting processes such that
 - P_0 is waiting for a resource that is held by P_1 ,
 - P_1 is waiting for a resource that is held by P_2 ,
 - ...,
 - P_{n-1} is waiting for a resource that is held by P_n , and
 - P_n is waiting for a resource that is held by P_0 .





Deadlock with Mutex Locks

- Deadlocks can occur via system calls, locking, etc
- See example box in text page 318 for mutex deadlock





Resource-Allocation Graph

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

- V is partitioned into two types:
 - $P = \{P_1, P_2, ..., P_n\}$, the set consisting of all the processes in the system
 - $R = \{R_1, R_2, ..., R_m\}$, the set consisting of all resource types in the system
- **request edge** directed edge $P_i \rightarrow R_j$
- **assignment edge** directed edge $R_j \rightarrow P_i$





Resource-Allocation Graph (Cont.)

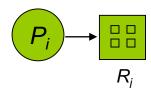
Process



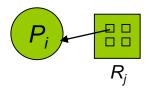
Resource Type with 4 instances



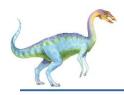
 \blacksquare P_i requests instance of R_j



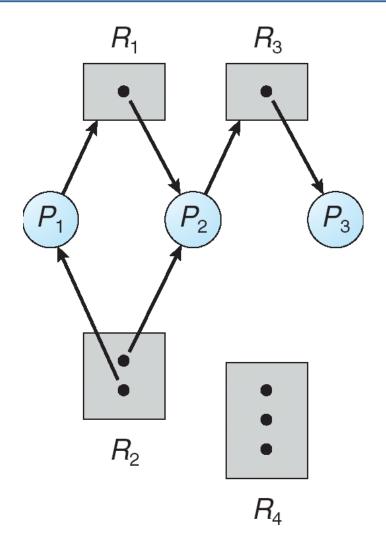
 \blacksquare P_i is holding an instance of R_j







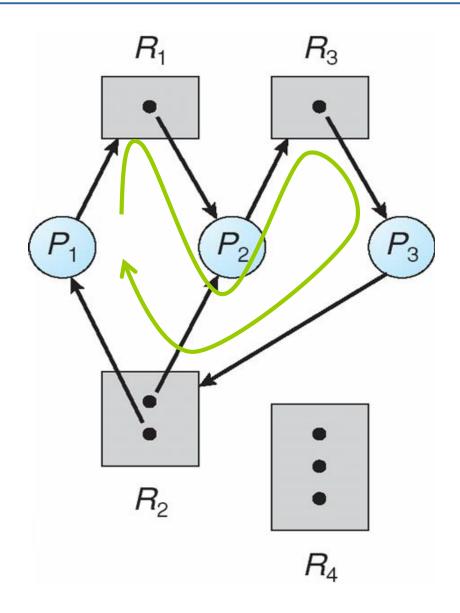
Example of a Resource Allocation Graph







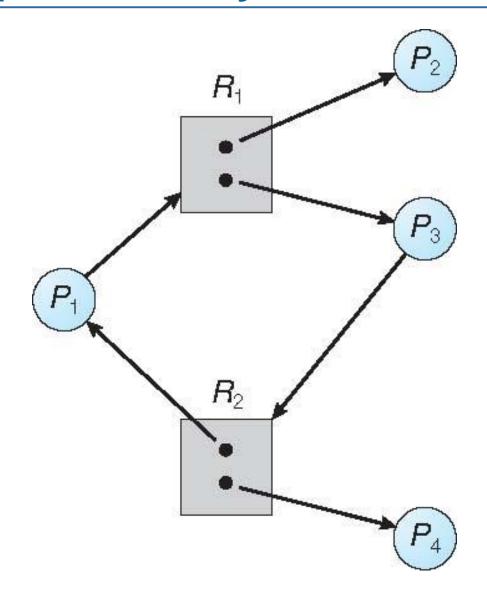
Resource Allocation Graph With A Deadlock



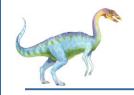




Graph With A Cycle But No Deadlock







Basic Facts

- If graph contains no cycles ⇒ no deadlock
- If graph contains a cycle ⇒
 - if only one instance per resource type, then deadlock
 - if several instances per resource type, possibility of deadlock





Methods for Handling Deadlocks

- Ensure that the system will never enter a deadlock state
- Allow the system to enter a deadlock state and then recover
- Ignore the problem and pretend that deadlocks never occur in the system;
 - Restart the system "manually" if the system "seems" to be deadlocked or stops functioning.
 - Note that the system may be "frozen" temporarily!
 - used by most operating systems, including UNIX





Deadlock Prevention

Restrain the ways request can be made

- Mutual Exclusion not required for sharable resources; must hold for nonsharable resources
- Hold and Wait must guarantee that whenever a process requests a resource, it does not hold any other resources
 - Require process to request and be allocated all its resources before it begins execution, or allow process to request resources only when the process has none
 - Low resource utilization; starvation possible





Deadlock Prevention (Cont.)

■ No Preemption –

- If a process that is holding some resources requests another resource that cannot be immediately allocated to it, then all resources currently being held are released
- Preempted resources are added to the list of resources for which the process is waiting
- Process will be restarted only when it can regain its old resources, as well as the new ones that it is requesting
- Circular Wait impose a total ordering of all resource types, and require that each process requests resources in an increasing order of enumeration



Deadlock Example

```
/* thread one runs in this function */
void *do work one(void *param)
  pthread mutex lock(&first mutex);
  pthread mutex lock(&second mutex);
   /** * Do some work */
  pthread mutex unlock(&second mutex);
   pthread mutex unlock(&first mutex);
  pthread exit(0);
/* thread two runs in this function */
void *do work two(void *param)
  pthread mutex lock(&second mutex);
  pthread mutex lock(&first mutex);
   /** * Do some work */
  pthread mutex unlock(&first mutex);
  pthread mutex unlock(&second mutex);
  pthread exit(0);
```



Deadlock Example with Lock Ordering

```
void transaction(Account from, Account to, double amount)
{
    mutex lock1, lock2;
    lock1 = get lock(from);
    lock2 = get lock(to);
    acquire(lock1);
        acquire(lock2);
        withdraw(from, amount);
        deposit(to, amount);
        release(lock2);
    release(lock1);
}
```





Deadlock Prevention

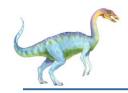
Circular Wait

A resource-ordering approach:

 $\begin{cases} F: R \rightarrow N \\ \text{Resource requests must be made in} \\ \text{an increasing order of enumeration.} \end{cases}$

- Type 1 strictly increasing order of resource requests.
 - Initially, order any # of instances of Ri
 - Following requests of any # of instances of Rj must satisfy
 F(Rj) > F(Ri), and so on.
 - * A single request must be issued for all needed instances of the same resources.





Deadlock Prevention

- Type 2
 - Processes must release all Ri's when they request any instance of Rj if F(Ri) ≥ F(Rj)
- F: R → N must be defined according to the <u>normal order</u> of resource usages in a system, e.g.,

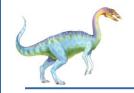
```
F(tape drive) = 1

F(disk drive) = 5

F(printer) = 12

?? feasible ??
```



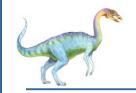


Deadlock Avoidance

Requires that the system has some additional *a priori* information available

- Simplest and most useful model requires that each process declare the maximum number of resources of each type that it may need
- The deadlock-avoidance algorithm dynamically examines the resource-allocation state to ensure that there can never be a circular-wait condition
- Resource-allocation state is defined by the number of available and allocated resources, and the maximum demands of the processes





Safe State

- When a process requests an available resource, system must decide if immediate allocation leaves the system in a safe state
- System is in safe state if there exists a sequence $\langle P_1, P_2, ..., P_n \rangle$ of ALL the processes in the systems such that for each P_i , the resources that P_i can still request can be satisfied by currently available resources + resources held by all the P_i , with i < 1

That is:

- If P_i resource needs are not immediately available, then P_i can wait until all P_j have finished
- When P_j is finished, P_i can obtain needed resources, execute, return allocated resources, and terminate
- When P_i terminates, P_{i+1} can obtain its needed resources, and so on



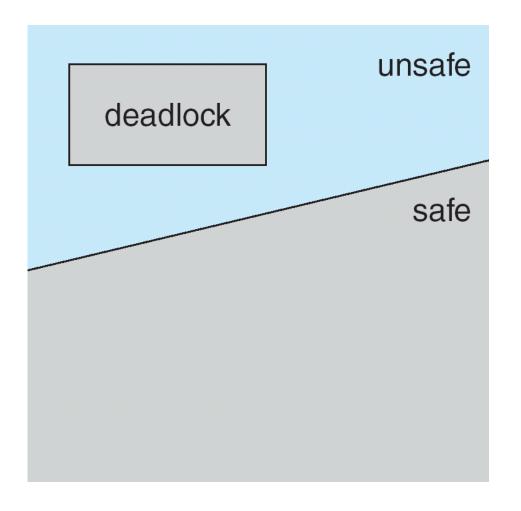
Basic Facts

- If a system is in safe state ⇒ no deadlocks
- If a system is in unsafe state ⇒ possibility of deadlock
- Avoidance ⇒ ensure that a system will never enter an unsafe state.

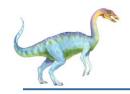




Safe, Unsafe, Deadlock State







Avoidance algorithms

- Single instance of a resource type
 - Use a resource-allocation graph
- Multiple instances of a resource type
 - Use the banker's algorithm





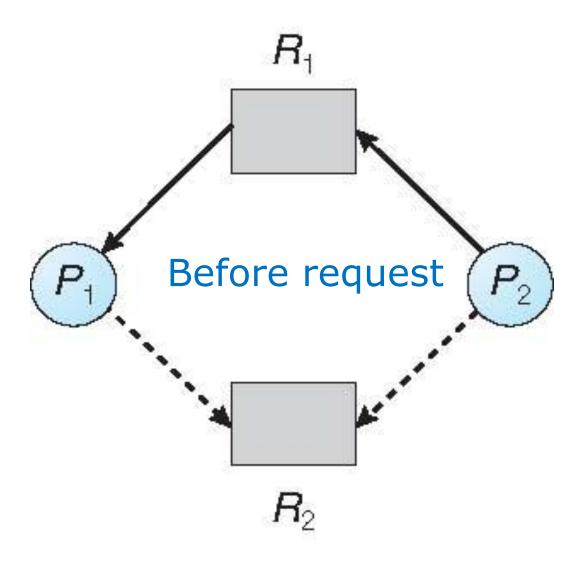
Resource-Allocation Graph Scheme

- Claim edge $P_i \rightarrow R_j$ indicated that process P_j may request resource R_i ; represented by a dashed line
- Claim edge converts to request edge when a process requests a resource
- Request edge converted to an assignment edge when the resource is allocated to the process
- When a resource is released by a process, assignment edge reconverts to a claim edge
- Resources must be claimed a priori in the system





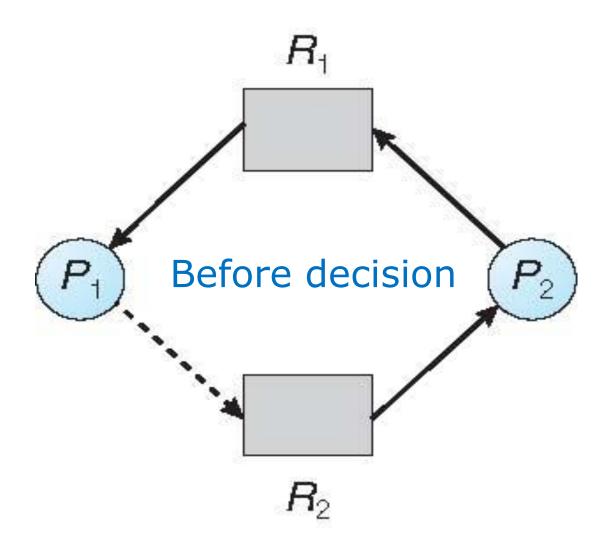
Resource-Allocation Graph







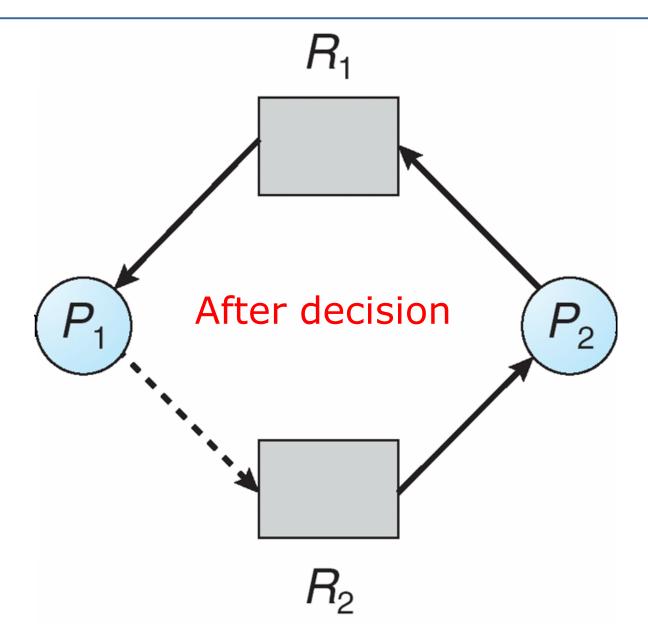
Unsafe State In Resource-Allocation Graph







Unsafe State In Resource-Allocation Graph

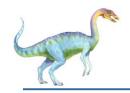


7.31



- Suppose that process P_i requests a resource R_i
- The request can be granted only if converting the request edge to an assignment edge does not result in the formation of a cycle in the resource allocation graph





Banker's Algorithm

- Multiple instances
- Each process must a priori claim maximum use
- When a process requests a resource it may have to wait
- When a process gets all its resources it must return them in a finite amount of time



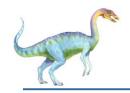
Data Structures for the Banker's Algorithm

Let n = number of processes, and m = number of resources types.

- Available: Vector of length m. If available [j] = k, there are k instances of resource type R_i available
- **Max**: $n \times m$ matrix. If Max[i,j] = k, then process P_i may request at most k instances of resource type R_i
- **Allocation**: $n \times m$ matrix. If Allocation[i,j] = k then P_i is currently allocated k instances of R_i
- **Need**: $n \times m$ matrix. If Need[i,j] = k, then P_i may need k more instances of R_i to complete its task

$$Need[i,j] = Max[i,j] - Allocation[i,j]$$





Safety Algorithm

1. Let **Work** and **Finish** be vectors of length *m* and *n*, respectively. Initialize:

Work = Available
Finish
$$[i]$$
 = false for $i = 0, 1, ..., n-1$

- 2. Find an *i* such that both:
 - (a) Finish [i] = false
 - (b) *Need_i* ≤ *Work*If no such *i* exists, go to step 4
- 3. Work = Work + Allocation; Finish[i] = true go to step 2

4. If Finish[i] == true for all i, then the system is in a safe state

Give students
5 mins so that
they can explian.

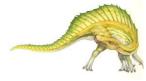
Resource-Request Algorithm for Process P_i

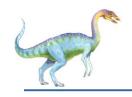
Request = request vector for process P_i . If **Request**_i[j] = k then process P_i wants k instances of resource type R_i

- If Request_i ≤ Need_i go to step 2. Otherwise, raise error condition, since process has exceeded its maximum claim
- 2. If $Request_i \le Available$, go to step 3. Otherwise P_i must wait, since resources are not available
- 3. Pretend to allocate requested resources to P_i by modifying the state as follows:

Available = Available - Request; Allocation; = Allocation; + Request;; Need; = Need; - Request;

- If safe ⇒ the resources are allocated to Pi
- If unsafe ⇒ Pi must wait, and the old resource-allocation state is restored





Example of Banker's Algorithm

• 5 processes P_0 through P_4 ;

3 resource types:

A (10 instances), B (5instances), and C (7 instances)

Snapshot at time T_0 :

	<u>Allocation</u>	<u>Max</u>	<u>Available</u>	<u>Need</u>
	ABC	ABC	ABC	ABC
P_0	010	753	332	7 4 3
P_1	200	322		122
P_2	302	902		600
P_3	211	222		011
P_4	002	4 3 3		4 3 1





Example (Cont.)

■ The content of the matrix *Need* is defined to be *Max* – *Allocation*

	<u>Need</u>	
	ABC	
P_0	7 4 3	
P_1	122	
P_2	600	
P_3	011	
P_4	4 3 1	

■ The system is in a safe state since the sequence $\langle P_1, P_3, P_4, P_2, P_0 \rangle$ satisfies safety criteria





Banker's algorithm Some issues

- How can we guess the right safe sequence?
- Can we always guess a right safe sequence?
- At step 2, when there is no i, why do not we need to backtrack?
 - If backtrack is needed, the banker's algorithm
 - collapses with huge complexity and
 - is no longer practical for run-time decision support.





Banker's algorithm

Example: P_1 Request (1,0,2)

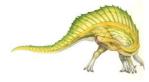
Check that Request \leq Available (that is, $(1,0,2) \leq (3,3,2) \Rightarrow$ true

	<u>Allocation</u>	<u>Need</u>	<u>Available</u>
	ABC	ABC	ABC
P_0	010	7 4 3	230
P_1	302	020	
P_2	3 0 1	600	
P_3	2 1 1	011	
P_4	002	4 3 1	

Executing safety algorithm shows that sequence

$$< P_1, P_3, P_4, P_0, P_2 >$$

satisfies safety requirement





Example: P_1 Request (1,0,2)

■ Check that Request \leq Available (that is, $(1,0,2) \leq (3,3,2) \Rightarrow$ true

	<u>Allocation</u>	<u>Need</u>	<u>Available</u>
	ABC	ABC	ABC
P_0	010	7 4 3	230
P_1	302	020	
P_2	302	600	
P_3	211	0 1 1	
P_4	002	431	

- Executing safety algorithm shows that sequence $< P_1, P_3, P_4, P_0, P_2 >$ satisfies safety requirement
- Can request for (3,3,0) by **P**₄ be granted?
- Can request for (0,2,0) by **P**₀ be granted?





Deadlock Detection

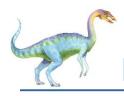
- Allow system to enter deadlock state
- Detection algorithm
- Recovery scheme



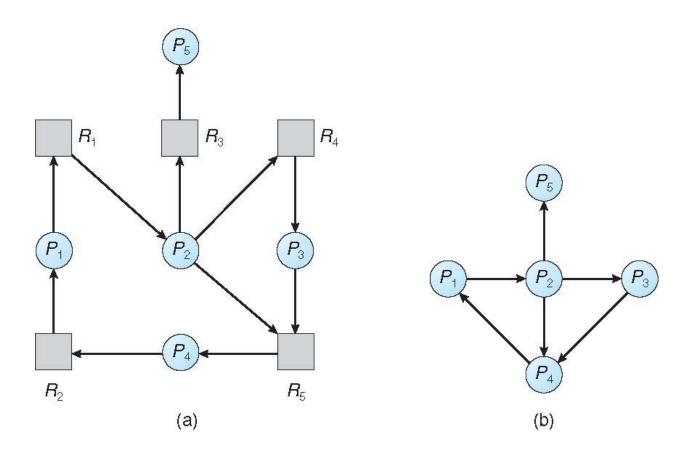


- Maintain wait-for graph
 - Nodes are processes
 - $P_i \rightarrow P_j$ if P_i is waiting for P_j
- Periodically invoke an algorithm that searches for a cycle in the graph. If there is a cycle, there exists a deadlock
- An algorithm to detect a cycle in a graph requires an order of n² operations, where n is the number of vertices in the graph





Resource-Allocation Graph and Wait-for Graph



Resource-Allocation Graph

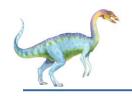
Corresponding wait-for graph





- Available: A vector of length m indicates the number of available resources of each type
- Allocation: An *n* x *m* matrix defines the number of resources of each type currently allocated to each process
- **Request**: An $n \times m$ matrix indicates the current request of each process. If Request[i][j] = k, then process P_i is requesting k more instances of resource type R_j .



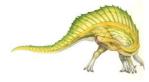


Detection Algorithm

- 1. Let **Work** and **Finish** be vectors of length **m** and **n**, respectively Initialize:
 - (a) Work = Available
 - (b) For i = 1,2, ..., n, if Allocation_i ≠ 0, then Finish[i] = false; otherwise, Finish[i] = true
- 2. Find an index *i* such that both:
 - (a) Finish[i] == false
 - (b) **Request**_i ≤ **Work**

If no such *i* exists, go to step 4

O(*m* x *n*²) operations





Detection Algorithm (Cont.)

- 3. Work = Work + Allocation; Finish[i] = true go to step 2
- 4. If *Finish[i]* == *false*, for some i, $1 \le i \le n$, then the system is in deadlock state. Moreover, if *Finish[i]* == *false*, then P_i is deadlocked

Algorithm requires an order of $O(m \times n^2)$ operations to detect whether the system is in deadlocked state





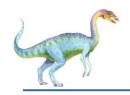
Example of Detection Algorithm

- Five processes P_0 through P_4 ; three resource types A (7 instances), B (2 instances), and C (6 instances)
- Snapshot at time T_0 :

	<u>Allocation</u>	<u>Request</u>	<u>Available</u>
	ABC	ABC	ABC
P_0	010	000	000
P_1	200	202	
P_2	303	000	
P_3	2 1 1	100	
P_4	002	002	

Sequence $\langle P_0, P_2, P_3, P_1, P_4 \rangle$ will result in **Finish[i] = true** for all **i**



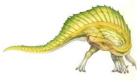


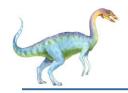
Example (Cont.)

P₂ requests an additional instance of type C

	<u>Reques</u>		
	ABC		
P_0	000		
P_1	202		
P_2	0 0 1		
P_3	100		
P_4	002		

- State of system?
 - Can reclaim resources held by process P_0 , but insufficient resources to fulfill other processes; requests
 - Deadlock exists, consisting of processes P₁, P₂, P₃, and P₄

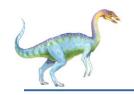




Detection-Algorithm Usage

- When, and how often, to invoke depends on:
 - How often a deadlock is likely to occur?
 - How many processes will need to be rolled back?
 - one for each disjoint cycle
- If detection algorithm is invoked arbitrarily, there may be many cycles in the resource graph and so we would not be able to tell which of the many deadlocked processes "caused" the deadlock.





Recovery from Deadlock: Process Termination

- Abort all deadlocked processes
- Abort one process at a time until the deadlock cycle is eliminated
- In which order should we choose to abort?
 - 1. Priority of the process
 - How long process has computed, and how much longer to completion
 - Resources the process has used
 - 4. Resources process needs to complete
 - 5. How many processes will need to be terminated
 - 6. Is process interactive or batch?





Recovery from Deadlock: Resource Preemption

- Selecting a victim minimize cost
- Rollback return to some safe state, restart process for that state
- Starvation same process may always be picked as victim, include number of rollback in cost factor





Exercise (1/2)

Exercises

Summary

A deadlocked state occurs when two or more processes are waiting indefinitely for an event that can be caused only by one of the waiting processes. There are three principal methods for dealing with deadlocks:

- · Use some protocol to prevent or avoid deadlocks, ensuring that the system will never enter a deadlocked state.
- Allow the system to enter a deadlocked state, detect it, and then recover.
- · Ignore the problem altogether and pretend that deadlocks never occur in the system.

The third solution is the one used by most operating systems, including Linux and Windows.

A deadlock can occur only if four necessary conditions hold simultaneously in the system: mutual exclusion, hold and wait, no preemption, and circular wait. To prevent deadlocks, we can ensure that at least one of the necessary conditions never holds.

A method for avoiding deadlocks, rather than preventing them, requires that the operating system have a priori information about how each process will utilize system resources. The banker's algorithm, for example, requires a priori information about the maximum number of each resource class that each process may request. Using this information, we can define a deadlockavoidance algorithm.

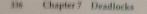
If a system does not employ a protocol to ensure that deadlocks will never occur, then a detection-and-recovery scheme may be employed. A deadlockdetection algorithm must be invoked to determine whether a deadlock has occurred. If a deadlock is detected, the system must recover either by terminating some of the deadlocked processes or by preempting resources from some of the deadlocked processes.

Where preemption is used to deal with deadlocks, three issues must be addressed: selecting a victim, rollback, and starvation. In a system that selects victims for rollback primarily on the basis of cost factors, starvation may occur, and the selected process can never complete its designated task.

Researchers have argued that none of the basic approaches alone is appropriate for the entire spectrum of resource-allocation problems in operating systems. The basic approaches can be combined, however, allowing us to select an optimal approach for each class of resources in a system.

Exercises

- 7.1 Consider the traffic deadlock depicted in Figure 7.10.
 - a. Show that the four necessary conditions for deadlock hold in this
 - b. State a simple rule for avoiding deadlocks in this system.



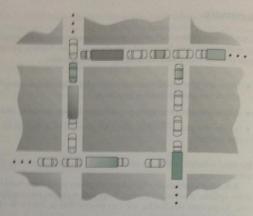


Figure 7.10 Traffic deadlock for Exercise 7.11,

- 7.2 Assume a multithreaded application uses only reader-writer locks in synchronization. Applying the four necessary conditions for deaded. is deadlock still possible if multiple reader-writer locks are used?
- 7.3 The program example shown in Figure 7.4 doesn't always lead to clearflock. Describe what role the CPU scheduler plays and how it can contribute to deadlock in this program.
- 7.4 In Section 7.4.4, we describe a situation in which we prevent deaded by ensuring that all locks are acquired in a certain order. However we also point out that deadlock is possible in this situation if the threads simultaneously invoke the transaction() function. Fa to transaction() function to prevent deadlocks.
- 7.5 Compare the circular-wait scheme with the various deadlock-avoidant schemes (like the banker's algorithm) with respect to the follows:
 - Runtime overheads
 - b. System throughput
- 7.6 In a real computer system, neither the resources available not demands of processes for resources are consistent over long per (months). Resources break or are replaced, new processes come and and new resources are bought and added to the system. If deadloo controlled by the banker's algorithm, which of the following characteristics can be made safely (without introducing the possibility of deadler and under what circumstances?
 - a. Increase Available (new resources added).





Exercise (2/2)

Exercises

- b. Decrease Available (resource permanently removed from system).
- Increase Max for one process (the process needs or wants more resources than allowed).
- d. Decrease Max for one process (the process decides it does not need that many resources).
- e. Increase the number of processes.
- Decrease the number of processes.
- 7.7 Consider a system consisting of four resources of the same type that are shared by three processes, each of which needs at most two resources. Show that the system is deadlock free.
- 7.8 Consider a system consisting of m resources of the same type being shared by n processes. A process can request or release only one resource at a time. Show that the system is deadlock free if the following two conditions hold:
 - The maximum need of each process is between one resource and m resources.
 - b. The sum of all maximum needs is less than m + n.
- 7.9 Consider the version of the dining-philosophers problem in which the chopsticks are placed at the center of the table and any two of them can be used by a philosopher. Assume that requests for chopsticks are made one at a time. Describe a simple rule for determining whether a particular request can be satisfied without causing deadlock given the current allocation of chopsticks to philosophers.
- 7.10 Consider again the setting in the preceding question. Assume now that each philosopher requires three chopsticks to eat. Resource requests are still issued one at a time. Describe some simple rules for determining whether a particular request can be satisfied without causing deadlock given the current allocation of chopsticks to philosophers.
- 7.11 We can obtain the banker's algorithm for a single resource type from the general banker's algorithm simply by reducing the dimensionality of the various arrays by 1. Show through an example that we cannot implement the multiple-resource-type banker's scheme by applying the single-resource-type scheme to each resource type individually.
- 7.12 Consider the following snapshot of a system:

	Allocation	Max
	ABCD	ABCD
Po	3014	5117
P_1	2210	3211
Po	3121	3321
P_3	0510	4612
P_4	4212	6325

338 Chapter 7 Deadlocks

Using the banker's algorithm, determine whether or not each of the following states is unsafe. If the state is safe, illustrate the order in which the processes may complete. Otherwise, illustrate why the state is unsafe.

a. Available =
$$(0, 3, 0, 1)$$

b.
$$Available = (1,0,0,2)$$

7.13 Consider the following snapshot of a system:

	Allocation	Max	Available
	ABCD	ABCD	ABCD
P_0	2001	4212	3321
Pı	3121	5252	
P ₂	2103	2316	
P_3	1312	1424	
P ₄	1432	3665	

Answer the following questions using the banker's algorithm:

- a. Illustrate that the system is in a safe state by demonstrating an order in which the processes may complete.
- b. If a request from process P_1 arrives for (1, 1, 0, 0), can the request be granted immediately?
- c. If a request from process *P*₄ arrives for (0, 0, 2, 0), can the request be granted immediately?
- 7.14 What is the optimistic assumption made in the deadlock-detection algorithm? How can this assumption be violated?
- 7.15 A single-lane bridge connects the two Vermont villages of North Tunbridge and South Tunbridge. Farmers in the two villages use this bridge to deliver their produce to the neighboring town. The bridge can become deadlocked if a northbound and a southbound farmer get on the bridge at the same time. (Vermont farmers are stubborn and are unable to back up.) Using semaphores and/or mutex locks, design an algorithm in pseudocode that prevents deadlock. Initially, do not be concerned about starvation (the situation in which northbound farmers prevent southbound farmers from using the bridge, or vice yersa).
- 7.16 Modify your solution to Exercise 7.15 so that it is starvation-free.

Programming Problems

7.17 Implement your solution to Exercise 7.15 using POSIX synchronization. In particular, represent northbound and southbound farmers as separate threads. Once a farmer is on the bridge, the associated thread will sleep for a random period of time, representing traveling across the bridge. Design your program so that you can create several threads representing the northbound and southbound farmers.

