Ch.5 – Process Synchronization

- Race Condition: several processes access and manipulate the same data concurrently, outcome depends on which order each access takes place.
- Each process has critical section of code, where it is manipulating data
  - To solve critical section problem each process must ask permission to enter critical section in entry section, follow critical section with exit section and then execute the remainder section
  - Especially difficult to solve this problem in preemptive kernels
- Peterson's Solution: solution for two processes
  - Two processes share two variables: `int` **`turn`** and `Boolean` **`flag[2]`**
  - **`turn:`** whose turn it is to enter the critical section
  - **`flag:`** indication of whether or not a process is ready to enter critical section
    - `flag[i] = true` indicates that process P_i is ready
  - Algorithm for process P_i:

```
do {
        flag[i] = TRUE;
        turn = j;
        while (flag[j] && turn == j)
                critical section
        flag[i] = FALSE;
        remainder section
} while (TRUE);
```

- Modern machines provide atomic hardware instructions: Atomic = non-interruptible
- Solution using Locks:

```
do {
        acquire lock
                critical section
        release lock
                remainder section
} while (TRUE);
```

- Solution using Test-And-Set: Shared boolean variable lock, initialized to FALSE

```
boolean TestAndSet (boolean *target){
        boolean rv = *target;
        *target = TRUE;"
        return rv:
}
```

```
do {
        while ( TestAndSet (&lock ))
                ; // do nothing
        // critical section
        lock = FALSE;
        // remainder section
} while (TRUE);
```

- Solution using Swap: Shared bool variable lock initialized to FALSE; Each process has local bool variable key

```
void Swap (boolean *a, boolean *b){
        boolean temp = *a;
        *a = *b;
        *b = temp:
}
```

```
do {
        key = TRUE;
        while ( key == TRUE)
                Swap (&lock, &key );
        // critical section
        lock = FALSE;
        // remainder section
} while (TRUE);
```

- Semaphore: Synchronization tool that does not require busy waiting (it has a busy-waiting version, by the non-busy-waiting version is the most common)
  - Standard operations: wait() and signal() ← these are the only operations that can access semaphore S
  - Can have counting (unrestricted range) and binary (0 or 1) semaphores
- Deadlock: Two or more processes are waiting indefinitely for an event that can be caused by only one of the waiting processes (most OSes do not prevent or deal with deadlocks)
  - Can cause starvation and priority inversion (lower priority process holds lock needed by higher-priority process)