Synchronization

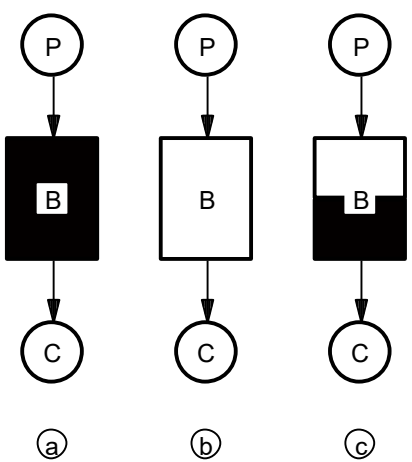
Classical Synchronization Problems

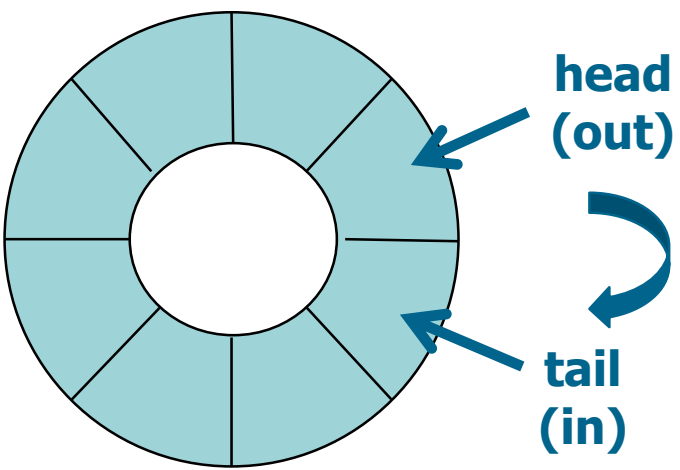
Stefano Quer and Stefano Scanzio
Dipartimento di Automatica e Informatica
Politecnico di Torino
skenz.it/os stefano.scanzio@polito.it

Producer-Consumer

- Producer and consumer with limited memory
  - It uses a circular buffer of dimension **SIZE** to store the elements to be produced and consumed
  - The circular buffer implements a FIFO queue (First-In First-Out)

full FIFO, empty, partially full





Sequential access

```
#define SIZE ...
...
int queue[SIZE];
int tail, head;
...
void init () {
  tail = 0;
  head = 0;
  n = 0;
}
```

FIFO standard (non ADT)

```
void enqueue (int val) {
  if (n>SIZE) return;
  queue[tail] = val;
  tail=(tail+1)%SIZE;
  n++;
  return;
}
```

```
void dequeue (int *val) {
  if (n<=0) return;
  *val=queue[head];
  head=(head+1)%SIZE;
  n--;
  return;
}
```

Sequential vs parallel access

- In the sequential access enqueue and dequeue are concurrent
- In parallel access we can have two cases
  - Only 1 producer and only 1 consumer
    - The operations enqueue and dequeue act on different extremes of the queue, however the n variable is shared
  - P producers and C consumers
    - In addition to the previous case, concurrent access operations to the same extreme of the queue are possible

Concurrent access: Version 1

- For parallel access with 1 producer and 1 consumer
  - You have to insert
    - A semaphore "full" that counts the number of filled elements
    - A semaphore "empty" that counts the number of empty elements
  - The counter n can be removed

Concurrent access: Version 1

```
#define SIZE ...
...
int queue[SIZE];
int tail, head;
...
void init () {
  tail = 0;
  head = 0;
}
```

FIFO standard (non ADT) without the variable n

```
void enqueue (int val) {
  queue[tail] = val;
  tail=(tail+1)%SIZE;
  return;
}
```

```
void dequeue (int *val) {
  *val=queue[head];
  head=(head+1)%SIZE;
  return;
}
```

Concurrent access: Version 1

1 Producer
1 Consumer

Instead of n it uses
elements filled
elements empty

```
init (full, 0);
init (empty, SIZE);
```

```
Producer () {
  int val;
  while (TRUE) {
    produce (&val);
    wait (empty);
    enqueue (val);
    signal (full);
  }
}
```

```
Consumer () {
  int val;
  while (TRUE) {
    wait (full);
    dequeue (&val);
    signal (empty);
    consume (val);
  }
}
```

Concurrent access: Version 2

- Solution 1 is simmetric
  - The producer produces filled positions
  - The consumer produces empty positions
- It can be easily **extended** to the case where there are more producers and more consumers
  - Producers and consumers operates on opposite extremes of the buffer
    - It can be done **concurrently**
    - As long as the queue is not completely full or completely empty
  - Instead, two producers or two consumers must act in **mutual exclusion**

Concurrent access: Version 2

P Producers
C Consumers

```
init (full, 0);
init (empty, SIZE);
init (MEp, 1);
init (MEc, 1);
```

It is necessary to force mutual exclusion between P and C

```
Producer () {
  int val;
  while (TRUE) {
    produce (&val);
    wait (empty);
    wait (MEp);
    enqueue (val);
    signal (MEp);
    signal (full);
  }
}
```

```
Consumer () {
  int val;
  while (TRUE) {
    wait (full);
    wait (MEc);
    dequeue (&val);
    signal (MEc);
    signal (empty);
    consume (val);
  }
}
```

Readers & Writers

- Classical problem
  - Courtois et al. [1971]
  - Share data between two sets of concurrent processes
    - A set of **Readers**, which can access **concurrently** to the data
    - A set of **Writers**, which can access in **mutual exclusion**, both with other Writers and Readers processes, to the data
  - Construct often used to create new synchronization primitives

Readers & Writers

- There are two versions of the problem
  - Precedence to Readers
  - Precedence to Writers
- Common goals
  - Respect the precedence protocol
  - Maximize concurrency

Precedence to Readers

- Giving precedence to Readers means
  - Privileging Readers access over Writers access, i.e.
  - Readers do not have to wait as long as a writer is in the CS
- Access protocol
  - Readers can concurrently access to the data
  - Until the Readers arrive, Writers have to wait
  - When even the last Reader ends, then you can wake up a writer (or a reader ... it depends on the scheduler)

Precedence to Readers: Version 1

```
nR = 0;
init (meR, 1);
init (w, 1);
```

Reader

```
wait (meR);
  nR++;
  if (nR==1)
    wait (w);
signal (meR);
...
reading
...
wait (meR);
  nR--;
  if (nR==0)
    signal (w);
signal (meR);
```

Writer

```
wait (w);
...
writing
...
signal (w);
```

Precedence to Readers: Version 2

```
nR = 0;
init (meR, 1);
init (meW, 1);
init (w, 1);
```

Reader

```
wait (meR);
  nR++;
  if (nR==1)
    wait (w);
signal (meR);
...
reading
...
wait (meR);
  nR--;
  if (nR==0)
    signal (w);
signal (meR);
```

To enforce the precedence to R (the signal(w) unblocks an R)

Writer

```
wait (meW);
wait (w);
...
writing
...
signal (w);
signal (meW);
```

Conclusions

- The solution uses
  - A global variable (nR) counts the number of Readers in the CS
  - A semaphore for the mutual exclusion for the access to the variable nR (meR)
  - A semaphore for the mutual exclusion of more Writers, or a Reader and the Writers (w)
  - Un semaforo di mutua eslusione per writer (meW)
- Writers are subject to **starvation**, because they can wait (be blocked) forever
- More complex solutions without starvation of the Writers are possible

Precedence to Writers

- Giving priority to writers means
  - A Writer that is ready, must wait the smallest possible time
- Access protocol
  - Each Writer must wait that all Readers finish
  - Each Writer has a higher priority than every Reader

Precedence to Writers

```
nR = nW = 0;
init (w, 1); init (r, 1);
init (meR, 1); init (meW, 1);
```

Reader

```
wait (r);
  wait (meR);
    nR++;
    if (nR == 1)
      wait (w);
  signal (meR);
signal (r);
...
reading
...
wait (meR);
  nR--;
  if (nR == 0)
    signal (w);
signal (meR);
```

Writer

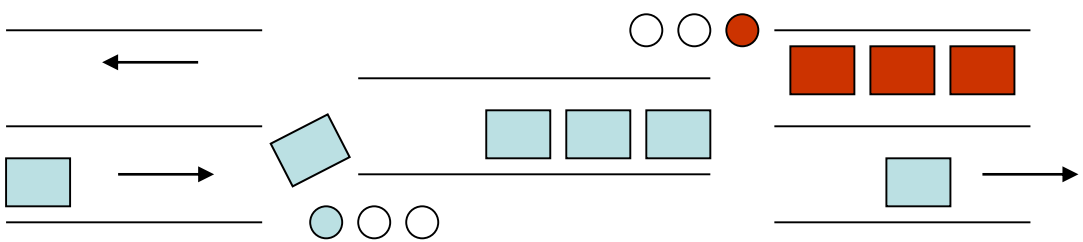
```
wait (meW);
  nW++;
  if (nW == 1)
    wait (r);
signal (meW);
wait (w);
  ...
  writing
  ...
signal (w)
wait (meW);
  nW--;
  if (nW == 0)
    signal (r);
signal (meW);
```

Conclusions

- The solution uses
  - Two global variables (nR and nW) to count the number of Readers and Writers
  - Two semaphores to guarantee mutual exclusion (meR and meW) for the access to the variables nR and nW
  - Two semaphores to guarantee mutual exclusion between Readers/Writers (r and w)
- Reader are subject to **starvation**, because they can wait (be blocked) forever
- More complex solutions without starvation are possible

The "Alternate direction tunnel"

- In an alternate direction tunnel
  - Allow any number of cars (processes) to proceed in the same direction
  - If there is traffic in one direction, block traffic in the opposite direction



The "Alternate direction tunnel"

- Extension to the Readers-Writers problem, with two sets of Readers
- Data structure
  - Two global counters (n1 and n2), one for each direction
  - Two semaphores (s1 and s2), one for each direction
  - A global semaphore for wait (busy)
- In its basic implementation, it can cause starvation of cars (in one direction with respect to the other)

Solution

```
n1 = n2 = 0;
init (s1, 1); init (s2, 1);
init (busy, 1);
```

left2right

```
wait (s1);
  n1++;
  if (n1 == 1)
    wait (busy);
signal (s1);
...
Run (left to right)
...
wait (s1);
  n1--;
  if (n1 == 0)
    signal (busy);
signal (s1);
```

right2left

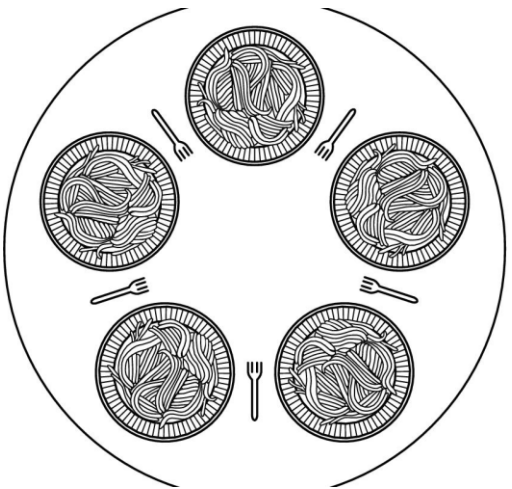
```
wait (s2);
  n2++;
  if (n2 == 1)
    wait (busy);
signal (s2);
...
Run (left to right)
...
wait (s2);
  n2--;
  if (n2 == 0)
    signal (busy);
signal (s2);
```

Dining (5) philosophers problem

- Model in which different resources are common to different concurrent processes
- Due to Dijkstra [1965]
- Definition of the problem
  - A table is set with
    - 5 rice dishes
    - 5 (Chinese) chopsticks each between two plates
  - Around the table sit 5 philosophers
  - Philosophers **think** or **eat**
    - To eat each philosopher needs two chopsticks
    - Chopsticks can be obtained one at a time

Model 0

- "Philosophical" solutions (not correct)
  - Teach philosophers to eat with only 1 chopstick
  - Provide more than 5 chopsticks
  - Allow only at most to 4 philosophers to sit at the table
  - Force asymmetry
    - Even position philosophers take the left fork first
    - Odd position philosophers take the right fork first

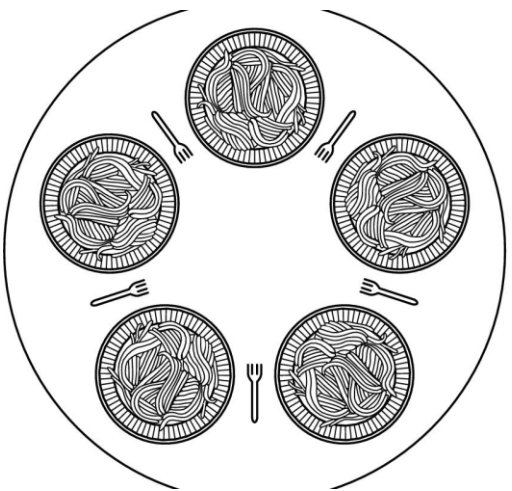


Model 1

- Use one binary semaphore (mutex) to protect the access to the only resource "the food"
  - Cancel concurrency
  - Only one philosopher eats at the same time (in two could eat)

```
init (mutex, 1);
```

```
while (true) {
   Think ();
   wait (mutex);
   Eat ();
   signal (mutex);
}
```



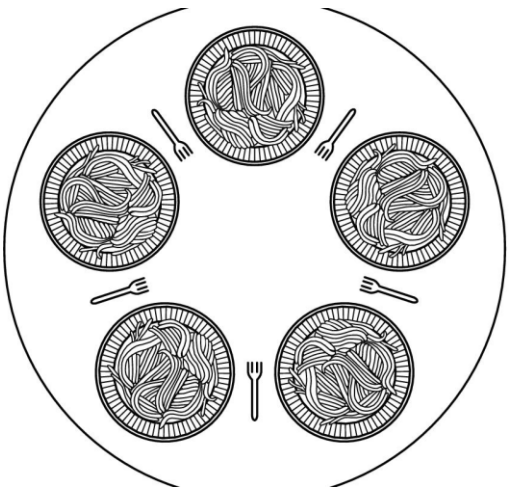
Model 2

- A semaphore for each chopstick
  - It can cause deadlock

```
init (chopstick[0], 1);
...
init (chopstick[4], 1);
```

```
while (true) {
  Think ();
  wait (chopstick[i]);
  wait (chopstick[(i+1)mod5]);
  Eat ();
  signal (chopstick[i]);
  signal (chopstick[(i+1)mod5]);
}
```

$i \in [0, 4]$



Solution

- Data structures
  - A state for each philosopher (THINKING, HUNGRY, EATING)
  - A semaphore for each philosopher (for access to food)
  - Another semaphore to manage the access in mutual exclusion to the philosopher state variable

```
while (TRUE) {
  Think ();
  takeForks (i);
  Eat ();
  putForks (i);
}
```

Solution

```
int state[N]
init (mutex, 1);
init (sem[0], 0); ...; init (sem[4], 0);
```

```
takeForks (int i) {
  wait (mutex);
  state[i] = HUNGRY;
  test (i);
  signal (mutex);
  wait (sem[i]);
}
```

```
putForks (int i) {
  wait (mutex);
  state[i] = THINKING;
  test (LEFT);
  test (RIGHT);
  signal (mutex);
}
```

```
test (int i) {
  if (state[i]==HUNGRY && state[LEFT]!=EATING &&
      state[RIGHT]!=EATING) {
    state[i] = EATING;
    signal (sem[i]);
  }
}
```