



Sistemas Distribuídos

Concurrency 2

António Rui Borges

Summary

A race condition is an undesirable situation that occurs when a device or system attempts to perform two or more operations at the same time, but because of the nature of the device or system, the operations must be done in the proper sequence to be done correctly.

- *General principles of concurrency*
 - *Critical regions*
 - *Racing conditions*
 - *Deadlock and indefinite postponement*
- *Synchronization devices*
 - *Monitors*
 - *Semaphores*
- *Java concurrency library*
- *Suggested readings*

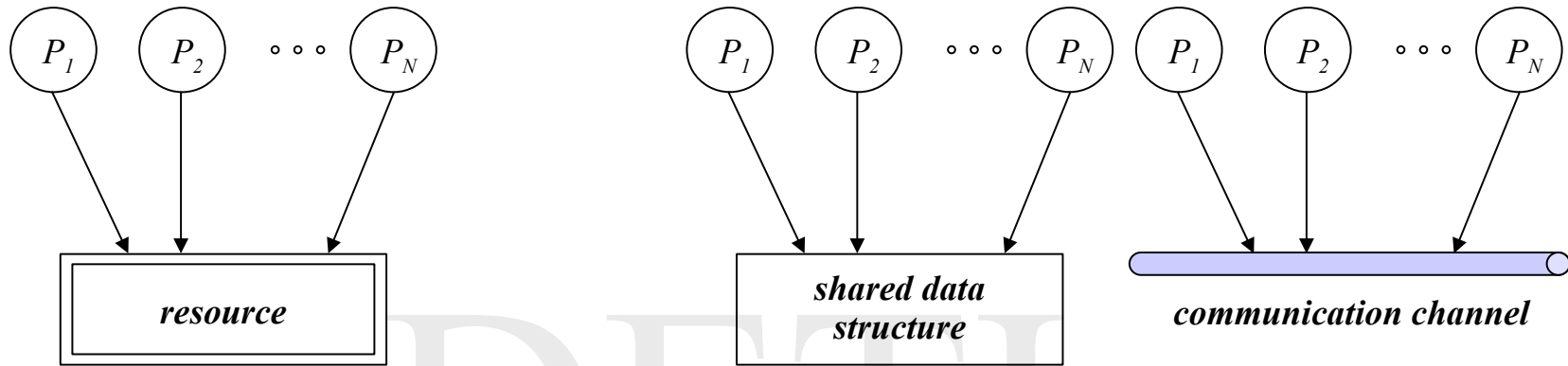
General principles of concurrency - 1

In a multiprogrammed environment, **coexisting processes** may present different behaviors interaction wise.

They may act as

- ***independent processes*** – when they are created, **live and die without explicitly interacting among themselves**; the underlying interaction is implicit and has its roots in the *competition* for the computational system resources; they are typically processes created by different users, or by the same user for different purposes, in an interactive environment, or processes which result from *job* processing in a *batch* environment
- ***cooperating processes*** – when they share information or communicate in an explicit manner; *sharing* presupposes a **common addressing space**, while *communication* can be carried out either by sharing the addressing space, or through a communication channel that connects the intervening processes.
a partilha tem de ser controlada pelo programador da app

General principles of concurrency - 2

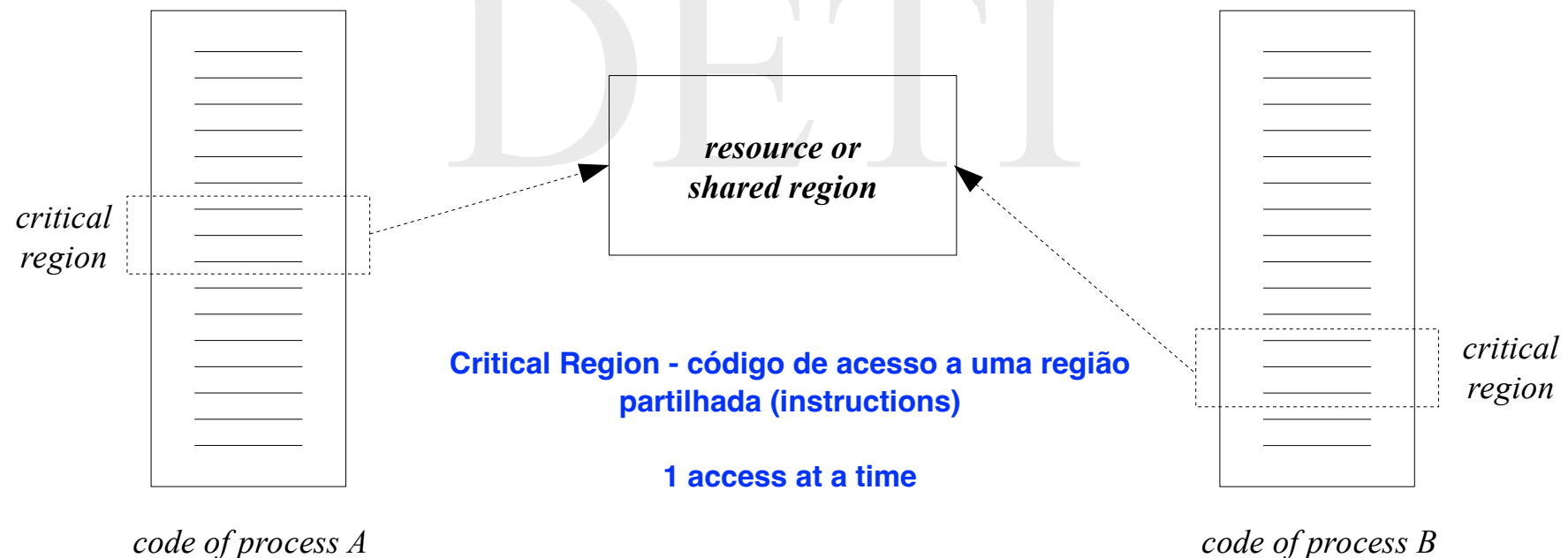


- *independent processes* that compete for access to a common resource of the computational system
- it is the *OS* responsibility to ensure that the resource assignment is carried out in a controlled fashion so that there is no loss of information
- this requires in general that only a single process may have access to the resource at a time (*mutual exclusion*)

- *cooperating processes* which share information or communicate among themselves
- it is the responsibility of the involved processes to ensure that access to the shared region is carried out in a controlled fashion so that there is no loss of information
- this requires in general that only a single process may have access at a time to the resource (*mutual exclusion*)
- the communication channel is typically a resource of the computational system; hence, access to it should be seen as *competition* for access to a common resource

General principles of concurrency - 3

Making the language precise, whenever one talks about *access by a process to a resource, or a shared region*, one is in reality talking about the processor executing the corresponding access code. This code, because it must be executed in a way that prevents the occurrence of *racing conditions*, which inevitably lead to loss of information, is usually called *critical region*.



General principles of concurrency - 4

livelock -> processos aguardam o acesso aos recursos no estado busy waiting, no deadlock estao no blocked

Imposing mutual exclusion on access to a resource, or to a shared region, can have, by its restrictive character, two undesirable consequences

- *deadlock / livelock* – it happens when two or more processes are waiting forever (blocked / in *busy waiting*) for the access to the respective critical regions, being held back by events which, one may prove, will never occur; as a result the operations can not proceed
- *indefinite postponement* – it happens when one or more processes compete for the access to a critical region and, due to a conjunction of circumstances where new processes come up continuously and compete with the former for this goal, access is successively denied; one is here, therefore, facing a real obstacle to their continuation.

One aims, when designing a multithreaded application, to prevent these pathological consequences to occur and to produce code which has a *liveness* property.

Problem of access to a critical region with mutual exclusion

Desirable properties that a general solution to the problem must assume

- *effective assurance of mutual exclusion imposition* – access to the critical region associated to a given resource, or shared region, can only be allowed to a single process at a time, among all that are competing to the access concurrently
- *independence on the relative speed of execution of the intervening processes, or of their number* – nothing should be presumed about these factors
- *a process outside the critical region can not prevent another to enter*
- *the possibility of access to the critical region of any process that wishes to can not be postponed indefinitely*
- *the time a process is inside a critical region is necessarily finite.*

Solution: Monitors and semaphores

Resources

Generally speaking, a *resource* is something a process needs to access. Resources may either be *physical components of the computational system* (processors, regions of the main or mass memory, specific input / output devices, etc), or *common data structures* defined at the operating system level (process control table, communication channels, etc) or among processes of an application.

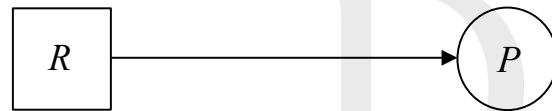
An essential property of resources is the kind of appropriation processes make of them. In this sense, resources are divided in

- *preemptable resources* – when they can be taken away from the processes that hold them, without any malfunction resulting from the fact; the processor and regions of the main memory where a process addressing space is stored, are examples of this class in multiprogrammed environments
- *non-preemptable resources* – when it is not possible; the printer or a shared data structure, requiring mutual exclusion for its manipulation, are examples of this class.

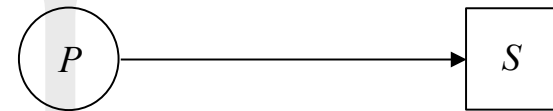
Schematic deadlock characterization

In a *deadlock* situation, only *non-preemptable* resources are relevant. The remaining can always be taken away, if necessary, from the processes that hold them and assigned to others to ensure that the latter may progress.

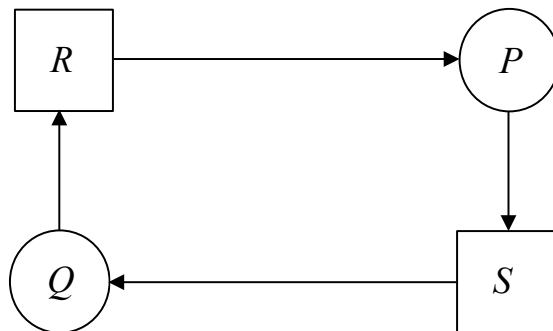
Thus, using this kind of framework, one may develop a schematic notation which represents a *deadlock* situation graphically.



process P holds the resource R



process P requires the resource S



typical deadlock situation
(the most simple one)

Necessary conditions to the occurrence of deadlock

It can be shown that, whenever *deadlock* occurs, there are four conditions which are necessarily present. They are the

- *condition of mutual exclusion* – each existing resource is either free, or was assigned to a single process (its holding can not be shared)
- *condition of waiting with retention* – each process, upon requesting a new resource, holds all other previously requested and assigned resources
- *condition of non-liberation* – nothing, but the process itself, can decide when a previously assigned resource is freed
- *condition of circular waiting (or vicious circle)* – a circular chain of processes and resources, where each process requests a resource which is being held by the next process in the chain, is formed.

Deadlock prevention - 1

The necessary conditions to the occurrence of *deadlock* lead to the statement

deadlock occurrence \Rightarrow *mutual exclusion on access to a resource* **and**
waiting with retention **and**
no liberation of resources **and**
circular waiting

which is equivalent to

no mutual exclusion on access to a resource **or**
no waiting with retention **or**
liberation of resources **or**
no circular waiting \Rightarrow *no deadlock occurrence* .

Thus, in order to make deadlock *impossible to happen*, one has only to deny one of the necessary conditions to the occurrence of deadlock. Policies which follow this strategy are called *deadlock prevention policies*.

Deadlock prevention - 2

The first, *mutual exclusion on access to a resource*, is too restrictive because it can only be denied for non-preemptable resources. Otherwise, *racing conditions* are introduced which lead, or may lead, to information inconsistency.

Reading access by multiple processes to a file is a typical example of denying this condition. One should point out that, in this case, it is also common to allow at the same time a single writing access. When this happens, however, *racing conditions*, with the consequent loss of information, can not be completely discarded. *Why?*

Therefore, only the last three conditions are usually object of denial.

Denying the condition of waiting with retention

It means that a *process must request at once all the resources it needs for continuation*. If it can get hold of them, the completion of the associated activity is ensured. Otherwise, it must wait.

One should notice that *indefinite postponement* is not precluded. The procedure must also ensure that sooner or later the necessary resources will always be assigned to any process which will be requesting them. The introduction of *aging* policies to increase the priority of a process is a very popular method used in this situation.

Imposing the condition of liberation of resources

It means that a *process*, when it can not get hold of all the resources it requires for continuation, must release all the resources in its possession and start later on the whole request procedure from the very beginning. Alternatively, it also means that a process can only hold a resource at a time (this, however, is a particular solution and is not applicable in most cases).

Care should be taken for the process not to enter a *busy waiting* procedure of request / acquire resources. In principle, the process must block after freeing the resources it holds and be waken up only when the resources it was requesting are released.

Nevertheless, *indefinite postponement* is not precluded. The procedure must also ensure that sooner or later the necessary resources will always be assigned to any process which will be requesting them. The introduction of *aging* policies to increase the priority of a process is a very popular method used in this situation.

Denying the condition of circular waiting

It means *to establish a linear ordering of the resources* and *to make the process, when it tries to get hold of the resources it needs for continuation, to request them in increasing order of the number associated to each of them.*

In this way, the possibility of formation of a circular chain of processes holding resources and requesting others is prevented.

One should notice that *indefinite postponement* is not precluded. The procedure must also ensure that sooner or later the necessary resources will always be assigned to any process which will be requesting them. The introduction of *aging* policies to increase the priority of a process is a very popular method used in this situation.

Monitors - 1

mutex -> mutual exclusion

A *monitor* is a synchronization device, proposed independently by Hoare and Brinch Hansen, which can be thought of as a special module defined within the [concurrent] programming language and consisting of an internal data structure, initialization code and a set of access primitives.

monitor example

```
(* internal data structure
   only accessible from the outside through the access primitives *)
var
  val: DATA;                                (* shared region *)
  c: condition;                             (* condition variable for synchronization *)
  (* access primitives *)
procedure pa1 (...);
end (* pa1 *)
function pa2 (...): real;
end (* pa2 *)
  (* initialization *)
begin
  ...
end
end monitor;
```


Monitors - 2

An application written in a concurrent language, implementing the *shared variables paradigm*, is seen as a set of *threads* that compete for access to shared data structures. When the data structures are implemented as *monitors*, the programming language ensures that the execution of a *monitor* primitive is carried out following a mutual exclusion discipline. Thus, the compiler, on processing a *monitor*, generates the required code to impose this condition in a manner totally transparent to the applications programmer.

A *thread* enters a *monitor* by calling one of its primitives, which constitutes the only way to access the internal data structure. As primitive execution entails mutual exclusion, when another *thread* is presently inside the monitor, the *thread* is blocked at the entrance, waiting for its turn.

Monitors - 3

Synchronization among *threads* using monitors is managed by *condition variables*. *Condition variables* are special devices, defined inside a monitor, where a thread may be blocked, while waiting for an event that allows its continuation to occur. There are two atomic operations which can be executed on a *condition variable*

wait – the calling *thread* is blocked at the *condition variable* passed as argument and is placed *outside the monitor* to allow another *thread*, wanting to enter, to proceed

signal – if there are blocked *threads* in the *condition variable* passed as argument, one of them is waken up; otherwise, nothing happens.

Monitors - 4

To prevent the coexistence of multiple *threads* inside a **monitor**, a rule is needed which states how the contention arisen by a *signal* execution is resolved

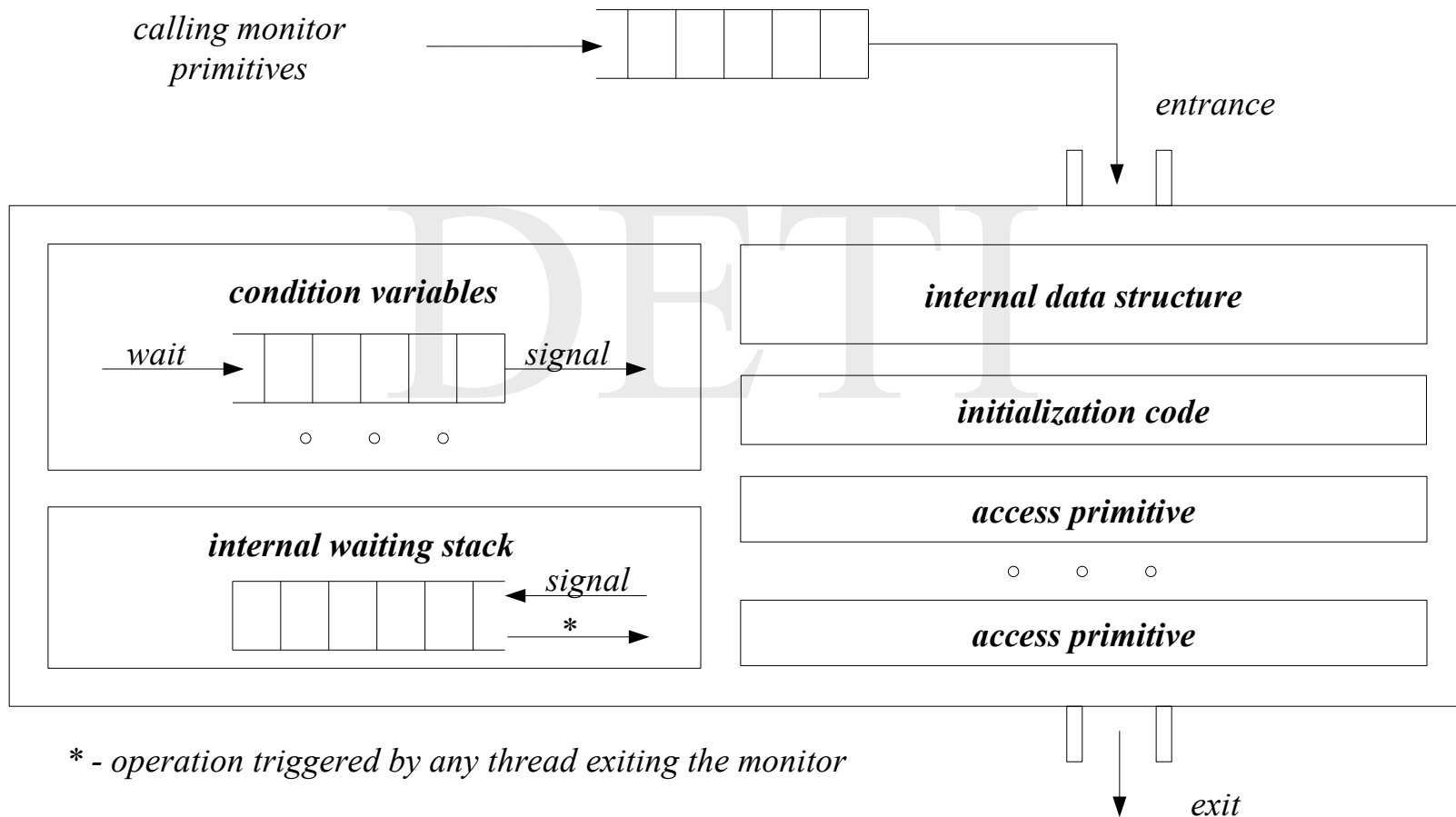
Hoare monitor – the *thread* which calls *signal* is placed *outside the monitor* so that the waken up *thread* may proceed; it is a very general solution, but its implementation requires a *stack*, where the *signal* calling *threads* are stored

Brinch Hansen monitor – the *thread* which calls *signal* must immediately exit the *monitor* (*signal* should be the very last executed instruction in any access primitive, except for a possible *return*); it is quite simple to implement, but it may become rather restrictive because it reduces the number of *signal* calls to one

Lampson / Redell monitor – the *thread* which calls *signal* proceeds, the waken up *thread* is kept *outside the monitor* and must compete for access to it again; it is still simple to implement, but it may give rise to *indefinite postponement* of some of the involved *threads*.

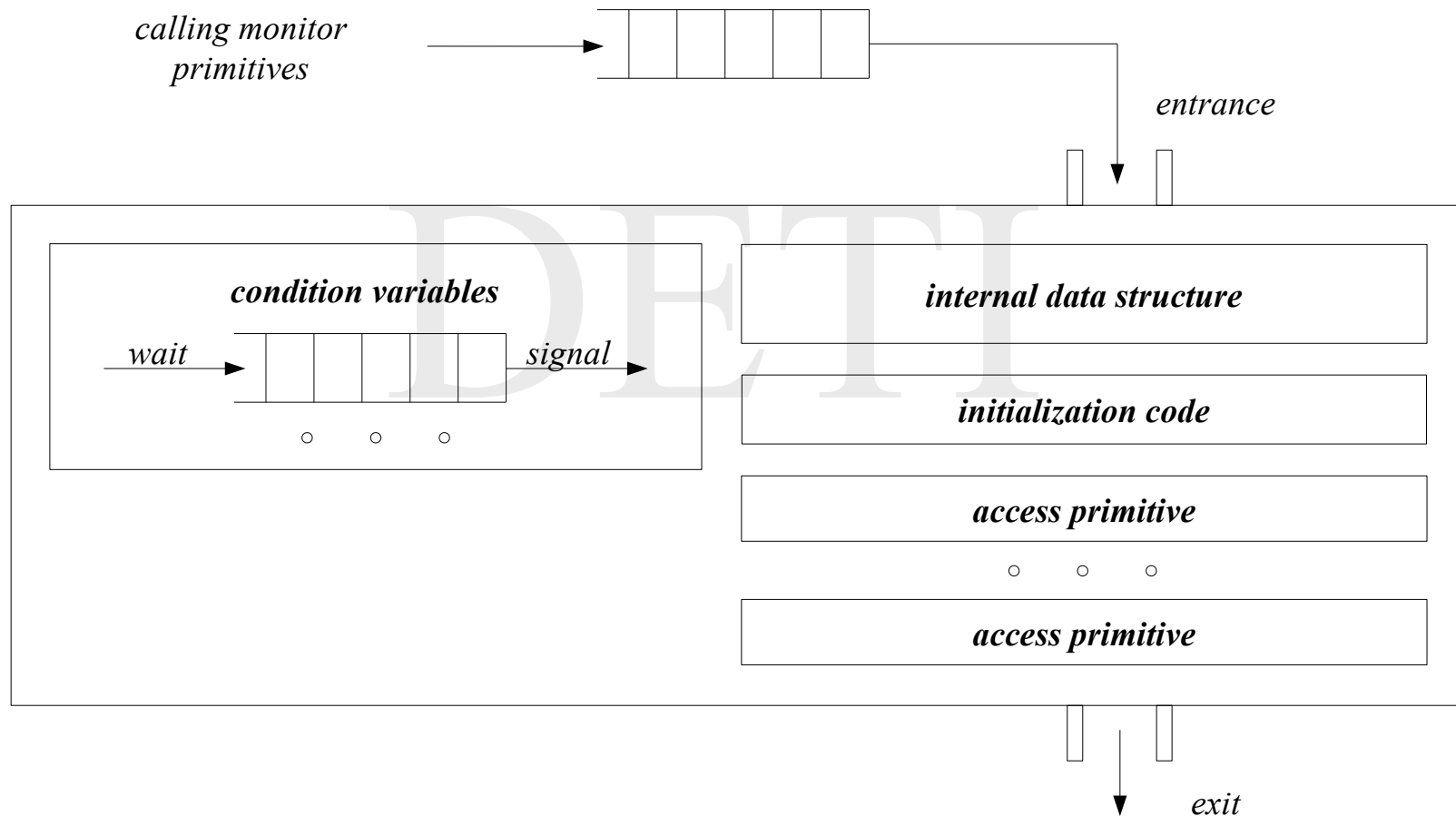
Monitors - 5

Hoare monitor



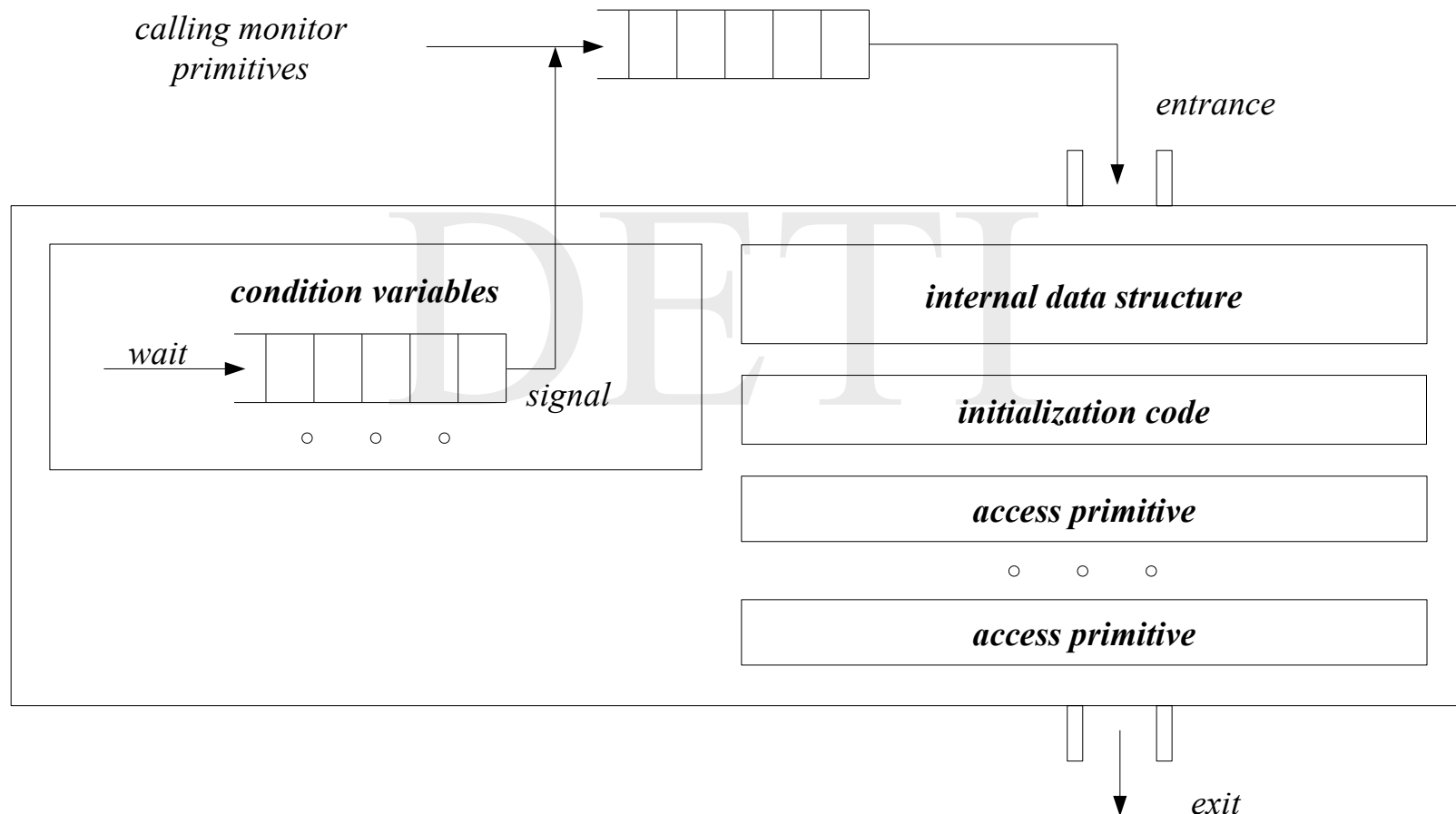
Monitors - 6

Brinch Hansen monitor



Monitors - 7

Lampson / Redell monitor



Monitors in Java - 1

Java supports Lampson / Redell monitors as its native synchronizing device

- each reference data type may be turned into a monitor, thus enabling to ensure mutual exclusion and *thread* synchronization when `static` methods are called upon it
- each instantiated object may be turned into a monitor, thus enabling to ensure mutual exclusion and *thread* synchronization when instantiation methods are called upon it.

In fact, and taking into account that Java is an object oriented language, each *thread*, being a Java object, can also be turned into a monitor. This property, if taken to the last consequences, enables a *thread* to block in its own monitor!

This, however, should never be done, since it introduces mechanisms of self-reference which are usually very difficult to understand due to the side effects they generate.

Monitors in Java - 2

The implementation in Java of a Lampson / Redell monitor has, however, some peculiarities

- the number of condition variables is limited to one, referenced in a implicit manner through the object which represents in *runtime* the reference data type, or any of its instantiations
- the traditional *signal* operation is named *notify* and there is a variant of this operation, *notifyAll*, the most commonly used, which enables the waking up of *all the threads* presently blocked in the condition variable
- furthermore, there is a method in data type `Thread`, named *interrupt*, which when called on a specific *thread*, aims to wake it up by throwing an *exception* if it is blocked on a condition variable.

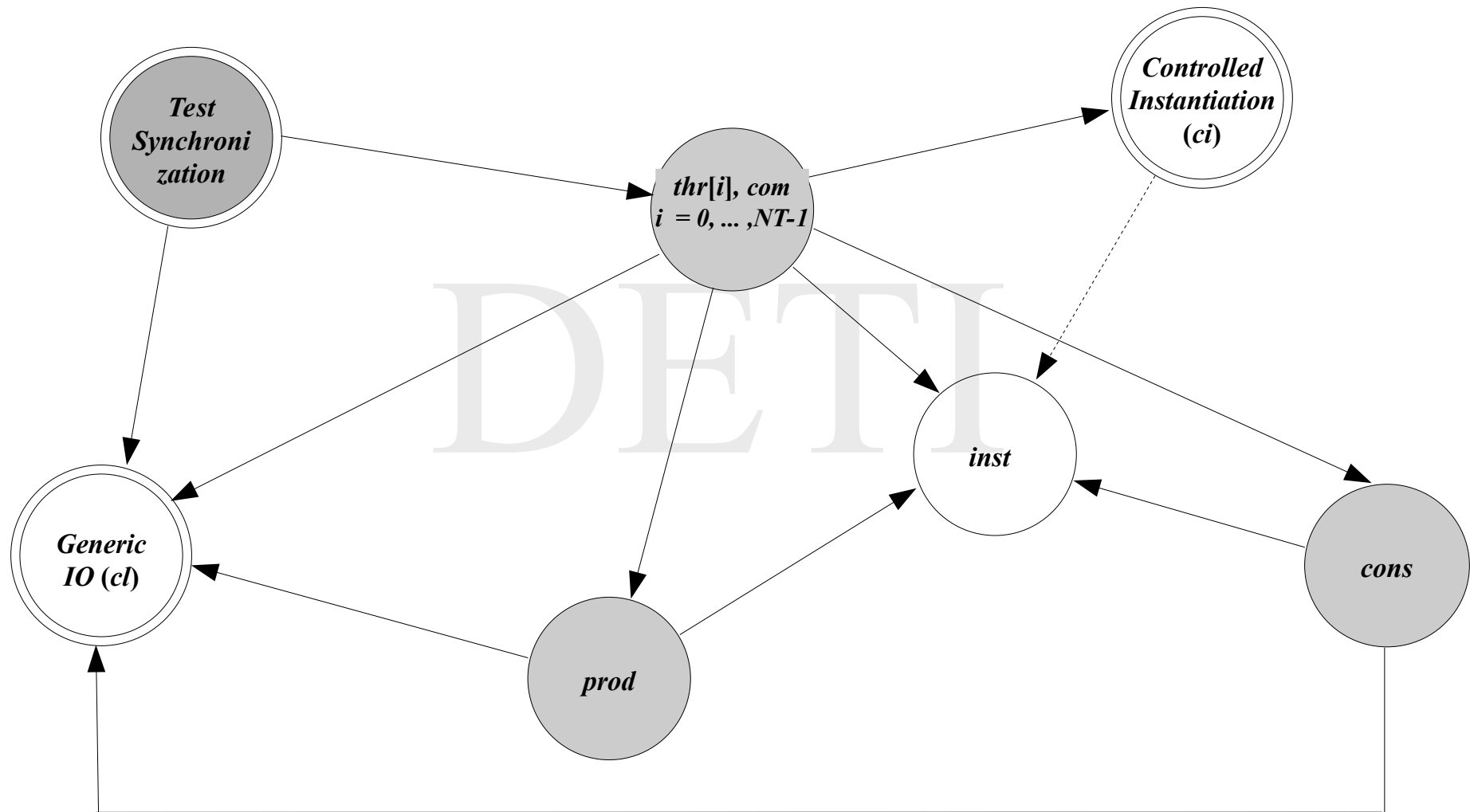
The need for the operation *notifyAll* becomes evident if one thinks that, by having a single condition variable per monitor, the only way to wake up a *thread* which is presently blocked waiting for a particular condition to be fulfilled, is by waking up *all* the blocked *threads* and then determine in a differential manner which is the one that can progress.

Controlled instantiation - 1

The example presented next illustrates the use of monitors in different situations

- *thread* `main` creates four *threads* of reference data type `TestThread` which try to instantiate objects of data type `ControlledInstantiation`, each a storage location for value transfer between *threads* of data types `Proc` and `Cons`
- the reference data type `ControlledInstantiation`, however, only allows the simultaneous instantiation of a maximum of two objects
- there are, thus, three types of monitors at play
 - the monitor associated with the object that represents in *runtime* the reference data type `ControlledInstantiation`
 - the monitor associated with each instantiation of data type `ControlledInstantiation`
 - the monitor associated with the object that represents in *runtime* the reference data type `genclass.GenericIO`
- the first controls the instantiation of `ControlledInstantiation` objects, the second the value transfer between threads of data type `Prod` and `Cons` and the third the printing of data in the *standard* output device.

Controlled instantiation - 2



Controlled instantiation - 3

Operation `wait` on a monitor defined at the reference data type level

```
public static synchronized ControlledInstantiation generateInst ()
{
    while (n >= NMAX)
    { try
      { ( Class.forName ("ControlledInstantiation")).wait ();
      }
      catch (ClassNotFoundException e)
      { GenericIO.writelnString ("Data type ControlledInstantiation was not " +
                                "found(generation)!");
        e.printStackTrace ();
        System.exit (1);
      }
      catch (InterruptedException e)
      { GenericIO.writelnString ("Static method generateInst was interrupted!");
      }
    }
    n += 1;
    nInst += 1;
    return new ControlledInstantiation ();
}
```

Controlled instantiation - 4

Operation notify on a monitor defined at the reference data type level

```
public static synchronized void releaseInst ()
{
    n -= 1;
    try
    { (Class.forName ("ControlledInstantiation")).notify ();
    }
    catch (ClassNotFoundException e)
    { GenericIO.writelnString ("Data type ControlledInstantiation was not found" +
                              "(release)!");
      e.printStackTrace ();
      System.exit (1);
    }
}
```

Controlled instantiation - 5

Operation `wait` and `notify` on a monitor defined at the object level

```
public synchronized void putVal (int val)
{
    store = val;
    while (store != -1)
    { notify ();
      try
      { wait ();
      }
      catch (InterruptedException e)
      { GenericIO.writelnString ("Method putVal was interrupted!");
      }
    }
}
```

DETI

Controlled instantiation - 6

Access with mutual exclusion to a non-thread safe library

```
Class<?> cl = null;           // representation of the library data type in JVM

try
{ cl = Class.forName ("genclass.GenericIO");
}
catch (ClassNotFoundException e)
{ System.out.println ("Data type genclass.GenericIO was not found!");
  e.printStackTrace ();
  System.exit (1);
}

synchronized (cl)
{ GenericIO.writelnString ("I have already created the threads!");
}
```

Controlled instantiation - 7

I have already created the threads!

I, Thread_base_2, got the instantiation number 2 of data type
ControlledInstantiation!

I, Thread_base_1, got the instantiation number 1 of data type
ControlledInstantiation!

I, Thread_base_1, am going to create the threads that will exchange the value!

I, Thread_base_2, am going to create the threads that will exchange the value!

I, Thread_base_1_writer, am going to write the value 1 in instantiation number 1
of data type ControlledInstantiation!

I, Thread_base_2_writer, am going to write the value 2 in instantiation number 2
of data type ControlledInstantiation!

I, Thread_base_1_reader, read the value 1 in instantiation number 1 of data type
ControlledInstantiation!

My thread which writes the value, Thread_base_1_writer, has terminated.

My thread which reads the value, Thread_base_1_reader, has terminated.

I, Thread_base_1, am going to release the instantiation number 1 of data type
ControlledInstantiation!

I, Thread_base_0, got the instantiation number 3 of data type
ControlledInstantiation!

I, Thread_base_0, am going to create the threads that will exchange the value!

I, Thread_base_2_reader, read the value 2 in instantiation number 2 of data type
ControlledInstantiation!

My thread which writes the value, Thread_base_2_writer, has terminated.

My thread which reads the value, Thread_base_2_reader, has terminated.

I, Thread_base_2, am going to release the instantiation number 2 of data type
ControlledInstantiation!

Controlled instantiation - 8

I, Thread_base_3, got the instantiation number 4 of data type ControlledInstantiation!
I, Thread_base_3, am going to create the threads that will exchange the value!
I, Thread_base_3_writer, am going to write the value 3 in instantiation number 4 of data type ControlledInstantiation!
I, Thread_base_0_writer, am going to write the value 0 in instantiation number 3 of data type ControlledInstantiation!
I, Thread_base_0_reader, read the value 0 in instantiation number 3 of data type ControlledInstantiation!
My thread which writes the value, Thread_base_0_writer, has terminated.
My thread which reads the value, Thread_base_0_reader, has terminated.
I, Thread_base_0, am going to release the instantiation number 3 of data type ControlledInstantiation!
The thread Thread_base_0 has terminated.
The thread Thread_base_1 has terminated.
The thread Thread_base_2 has terminated.
I, Thread_base_3_reader, read the value 3 in instantiation number 4 of data type ControlledInstantiation!
My thread which writes the value, Thread_base_3_writer, has terminated.
My thread which reads the value, Thread_base_3_reader, has terminated.
I, Thread_base_3, am going to release the instantiation number 4 of data type ControlledInstantiation!
The thread Thread_base_3 has terminated.

Semaphores - 1

A different approach to solve the problem of access with mutual exclusion to a critical region is based on the observation.

```
/* control data structure */
#define R    ...    /* number of processes wishing to access a critical region,
                        pid = 0, 1, ..., R-1 */

shared unsigned int access = 1;

/* primitive for entering the critical region */
void enter_RC (unsigned int own_pid)
{
    if (access == 0) sleep (own_pid);
    else access -= 1;
}

/* primitive for exiting the critical region */
void exit_RC ()
{
    if (there are blocked processes) wake_up_one ();
    else access += 1;
}
```

Diagram illustrating the atomic operations in the semaphore code:

- The `enter_RC` function contains an atomic operation: `if (access == 0) sleep (own_pid); else access -= 1;`. This operation is highlighted in a box and labeled as **atomic operation** (its execution can not be interrupted).
- The `exit_RC` function contains an atomic operation: `if (there are blocked processes) wake_up_one (); else access += 1;`. This operation is highlighted in a box and labeled as **atomic operation** (its execution can not be interrupted).

Semaphores - 2

A *semaphore* is a synchronization device, originally invented by Dijkstra, which can be thought of as a variable of the type

```
typedef struct
{ unsigned int val; /* non-negative counting value */
  NODE *queue; /* blocked processes waiting queue */
} SEMAPHORE;
```

where two atomic operations may be performed

down – if `val` is different from zero (the semaphore has a shade of green), its value is decreased by one unit; otherwise, the process which called the operation, is blocked and its id is placed in `queue`

up – if there are blocked processes in `queue`, one of them is waken up (according to any previously prescribed discipline); otherwise, the value of `val` is increased by one unit.

Semaphores - 3

```
/* kernel defined semaphore array */
#define R    ...          /* semaphore id - id = 0, 1, ..., R-1 */

static SEMAPHORE sem[R];

/* down operation */
void down (unsigned int id)
{
    disable interrupts / lock access flag;
    if (sem[id].val == 0) sleep_on_sem (getpid(), id);
    else sem[id].val -= 1;
    enable interrupts / unlock access flag;
}

/* up operation */
void up (unsigned int id)
{
    disable interrupts / lock access flag;
    if (there are blocked processes in sem[id]) wake_up_one_on_sem (id);
    else sem[id].val += 1;
    enable interrupts / unlock access flag;
}
```

Java Dijkstra semaphore

```
public class Semaphore
{
    private int val = 0,                // green / red indicator
               numbBlockThreads = 0;    // number of the blocked threads
                                       // in the monitor

    public synchronized void down ()
    {
        if (val == 0)
        { numbBlockThreads += 1;
          try
          { wait ();
            }
          catch (InterruptedException e) {}
        }
        else val -= 1;
    }

    public synchronized void up ()
    {
        if (numbBlockThreads != 0)
        { numbBlockThreads -= 1;
          notify ();
        }
        else val += 1;
    }
}
```

Java concurrency library - 1

Java concurrency library supplies the following relevant synchronization devices

- *barriers* – devices which lead to the blocking of set of *threads* until a condition for their continuation is fulfilled; when the barrier is raised, all the blocked processes are waken up
- *semaphores* – devices that provide a more general implementation of the semaphore concept than the one prescribed by Dijkstra, namely allowing the calling of *down* and *up* operations where the internal field `val` is decreased or increased by more than an unit at a time and having a non-blocking variant of *down* operation
- *exchanger* – devices that allow an exchange of values between *threads* pairs through a *rendez-vous* type synchronization.

Java concurrency library - 2

- *locks* – Lampson-Redell monitors of general character (in contrast to Java *built-in* monitor, it is possible to define here multiple condition variables and to get a similar functionality to that provided by the *pthread* library)
- *atomic variable manipulation* – *read-modify-write* mechanisms which allow writing and reading the contents of several types of variables without the risk of *racing conditions*.

The library also provides basic thread blocking primitives for creating locks and other synchronization devices.

Suggested reading

- *Distributed Systems: Concepts and Design, 4th Edition*, Coulouris, Dollimore, Kindberg, Addison-Wesley
 - Chapter 6: *Operating systems support*
- *Distributed Systems: Principles and Paradigms, 2nd Edition*, Tanenbaum, van Steen, Pearson Education Inc.
 - Chapter 3: *Processes*
- *On-line* support documentation for Java program developing environment by Oracle (Java Platform Standard Edition 8)