3. Safety and Synchronization

Oscar Nierstrasz

Selected material © Magee and Kramer
Roadmap

> Modelling interaction in FSP
> Safety — synchronizing critical sections
> Locking for atomicity
> The busy-wait mutual exclusion protocol
> Checking Safety properties
> Conditional synchronization
Roadmap

> Modelling interaction in FSP
> Safety — synchronizing critical sections
> Locking for atomicity
> The busy-wait mutual exclusion protocol
> Checking Safety properties
> Conditional synchronization
Modelling interaction — shared actions

Actions that are common between two processes are **shared** and can be used to model **process interaction**:

> Unshared actions may be **arbitrarily interleaved**
> Shared actions occur **simultaneously** for all participants

```
MAKER  = ( make -> ready -> MAKER ).
USER    = ( ready -> use -> USER ).
MAKER_USER = ( MAKER || USER ).
```

**What are the states of the LTS?**

**The traces?**
Modelling interaction — handshake

A **handshake** is an action that signals acknowledgement.

\[
\text{MAKERv2} = ( \text{make} \rightarrow \text{ready} \rightarrow \text{used} \rightarrow \text{MAKERv2} ).
\]

\[
\text{USERv2} = ( \text{ready} \rightarrow \text{use} \rightarrow \text{used} \rightarrow \text{USERv2} ).
\]

\[
||\text{MAKER\_USERv2} = ( \text{MAKERv2} || \text{USERv2} ).
\]

**What are the states and traces of the LTS?**
Roadmap

- Modelling interaction in FSP
- **Safety — synchronizing critical sections**
- Locking for atomicity
- The busy-wait mutual exclusion protocol
- Checking Safety properties
- Conditional synchronization
Objects must only be accessed when they are in a **consistent state**, formalized by a **class invariant**.

Each method **assumes** the class invariant holds when it starts, and it **re-establishes** it when done.

If methods interleave arbitrarily, an **inconsistent state** may be accessed, and the object may be left in a “dirty” state.
Race conditions

A race condition exists if safety may be violated by bad timing

```java
public class AccountBAD extends Account {
    // unsynchronized!
    public void withdraw(int amount) {
        while (amount > assets) {
            Thread.yield();  // busy wait
        }
        Thread.yield();  // race condition!
        assets -= amount;
        checkInvariant();  // might fail!
    }
}
```

```java
if (!assets >= 0) { errors++; }
```

Consider the two processes:

\[
\begin{align*}
\{ & x = 0 \} \\
\text{AInc: } & x := x + 1 \\
\text{BInc: } & x := x + 1 \\
\{ & x = ? \}
\end{align*}
\]

How can these processes interfere?
Atomic actions

Individual reads and writes may be atomic actions:

\[
\begin{align*}
\text{const} & \quad N = 3 \\
\text{range} & \quad T = 0..N \\
\text{Var} & \quad = \, \text{Var}[0], \\
\text{Var}[u:T] & \quad = \ ( \text{read}[u] \rightarrow \text{Var}[u] \\
& \quad \mid \text{write}[v:T] \rightarrow \text{Var}[v] ) . \\
\text{set} & \quad \text{VarAlpha} = \{ \text{read}[T], \text{write}[T] \} \\
\text{Inc} & \quad = \ ( \text{read}[v:0..N-1] \\
& \quad \rightarrow \text{write}[v+1] \\
& \quad \rightarrow \text{STOP } ) +\text{VarAlpha}.
\end{align*}
\]
Sequential behaviour

A single sequential thread requires no synchronization:

\[
\text{SeqInc} = (\text{Var} \parallel \text{Inc})
\]
Roadmap

> Modelling interaction in FSP
> Safety — synchronizing critical sections
> **Locking for atomicity**
> The busy-wait mutual exclusion protocol
> Checking Safety properties
> Conditional synchronization
Without synchronization, concurrent threads may interfere:

\[
\text{ParInc} = (\{a,b\}::\text{Var} \mid \mid a::\text{Inc} \mid \mid b::\text{Inc}).
\]
Locking

Locks are used to make a critical section atomic:

\[
\text{LOCK} = ( \text{acquire} \rightarrow \text{release} \rightarrow \text{LOCK} ).
\]

\[
\text{INC} = ( \text{acquire} \\
\rightarrow \text{read}[v:0..N-1] \\
\rightarrow \text{write}[v+1] \\
\rightarrow \text{release} \\
\rightarrow \text{STOP} ) + \text{VarAlpha}.
\]
Processes can synchronize critical sections by sharing a lock:

\[ \text{ParInc2} = (\{a,b\}::\text{VAR} \mid \{a,b\}::\text{LOCK} \mid a::\text{INC} \mid b::\text{INC}). \]
Synchronization in Java

Java Threads also synchronize using locks:

```java
synchronized T m() {
    // method body
}
```

is just *convenient syntax* for:

```java
T m() {
    synchronized (this) {
        // method body
    }
}
```

*Every object has a lock*, and Threads may use them to synchronize with each other.
Roadmap

> Modelling interaction in FSP
> Safety — synchronizing critical sections
> Locking for atomicity
> **The busy-wait mutual exclusion protocol**
> Checking Safety properties
> Conditional synchronization
Busy-Wait Mutual Exclusion Protocol

P1 sets \( \text{enter1} := \text{true} \) when it wants to enter its CS, but sets \( \text{turn} := \text{"P2"} \) to yield priority to P2.

process P1
  loop
    enter1 := true
    turn := "P2"
    while enter2 and turn = "P2"
      do skip
        Critical Section
    enter1 := false
    Non-critical Section
  end
end

process P2
  loop
    enter2 := true
    turn := "P1"
    while enter1 and turn = "P1"
      do skip
        Critical Section
    enter2 := false
    Non-critical Section
  end
end

Is this protocol correct? Is it fair? Deadlock-free?
Atomic read and write

We can model integer and boolean variables as processes with atomic read and write actions:

\[
\text{range } T = 1..2 \\
\text{Var} = \text{Var}[1], \\
\text{Var}[u:T] = \\
\quad (\text{read}[u] \rightarrow \text{Var}[u] \\
\quad | \text{write}[v:T] \rightarrow \text{Var}[v]).
\]

\[
\text{set Bool} = \{\text{true, false}\} \\
\text{BOOL(Init='false)} = \text{BOOL}[\text{Init}], \\
\text{BOOL}[b:\text{Bool}] = \\
\quad (\text{is}[b] \rightarrow \text{BOOL}[b] \\
\quad | \text{setTo}[x:\text{Bool}] \rightarrow \text{BOOL}[x]).
\]
Modelling the busy-wait protocol

Each process performs two actions in its CS:

\[ P_1 = ( \text{enter1}.\text{setTo}['true'] \]
\[
\quad \rightarrow \text{turn}.\text{write}[2]
\]
\[
\quad \rightarrow \text{Gd1},
\]
\[ \text{Gd1} = \]
\[
\big( \text{enter2}.\text{is}['false'] \rightarrow \text{CS1} \big)
\]
\[
\quad \big| \text{enter2}.\text{is}['true'] \rightarrow \]
\[
\big( \text{turn}.\text{read}[1] \rightarrow \text{CS1} \big)
\]
\[
\big| \text{turn}.\text{read}[2] \rightarrow \text{Gd1}),
\]
\[ \text{CS1} = ( a \rightarrow b 
\]
\[
\quad \rightarrow \text{enter1}.\text{setTo}['false']
\]
\[
\quad \rightarrow P_1).\]

\[ P_2 = ( \text{enter2}.\text{setTo}['true'] \]
\[
\quad \rightarrow \text{turn}.\text{write}[1]
\]
\[
\quad \rightarrow \text{Gd2},
\]
\[ \text{Gd2} = \]
\[
\big( \text{enter1}.\text{is}['false'] \rightarrow \text{CS2} \big)
\]
\[
\quad \big| \text{enter1}.\text{is}['true'] \rightarrow \]
\[
\big( \text{turn}.\text{read}[2] \rightarrow \text{CS2} \big)
\]
\[
\big| \text{turn}.\text{read}[1] \rightarrow \text{Gd2}),
\]
\[ \text{CS2} = ( c \rightarrow d 
\]
\[
\quad \rightarrow \text{enter2}.\text{setTo}['false']
\]
\[
\quad \rightarrow \text{P2}).\]

\[ |\text{BusyWait} = (\text{enter1:BOOL} | \text{enter2:BOOL} | \text{turn:Var} | \text{P1} | \text{P2})@\{a,b,c,d\}.\]
Busy-wait composition

Very pretty, but how do we know there are no errors?!
We can check for errors by composing our system with an agent that moves to the ERROR state if atomicity is violated:

\[
\text{Ok} = (\ a \rightarrow (\ c \rightarrow \text{ERROR} \ |\ b \rightarrow \text{Ok}) \\
|\ c \rightarrow (\ a \rightarrow \text{ERROR} \ |\ d \rightarrow \text{Ok})).
\]

|BusyWaitOk = (\text{enter1:BOOL}|\text{enter2:BOOL}|\text{turn:Var}|\text{P1}|\text{P2}|\text{Ok}).
Roadmap

- Modelling interaction in FSP
- Safety — synchronizing critical sections
- Locking for atomicity
- The busy-wait mutual exclusion protocol
- Checking Safety properties
- Conditional synchronization
A safety property asserts that nothing bad happens. Use the ERROR process (-1) to detect erroneous behaviour.

\[
\text{ACTUATOR} = (\text{command} \rightarrow \text{ACTION}),
\]
\[
\text{ACTION} = (\text{respond} \rightarrow \text{ACTUATOR} | \text{command} \rightarrow \text{ERROR}).
\]

Trace to property violation in ACTUATOR:
command
command
Safety — property specification

ERROR conditions state what is not required
In complex systems, it is usually better to specify directly what is required.

property SAFE_ACTUATOR
= ( command
  -> respond
  -> SAFE_ACTUATOR ) .

Trace to property violation in SAFE_ACTUATOR:
respond
Checking Busy-Wait (revisited)

property \( \text{Ok} = \) ( \( a \rightarrow b \rightarrow \text{Ok} \) \\
| \( c \rightarrow d \rightarrow \text{Ok} \) ).

Ok is a \textit{deterministic process} that specifies which traces are safe — everything else is in ERROR.

\textbf{Contrast:}

\( \text{Ok} = \) ( \( a \rightarrow ( c \rightarrow \text{ERROR} \mid b \rightarrow \text{Ok} ) \) \\
| \( c \rightarrow ( a \rightarrow \text{ERROR} \mid d \rightarrow \text{Ok} ) \) ).
A **safety property** $P$ defines a *deterministic process* that asserts that any trace including actions in the alphabet of $P$ is accepted by $P$.

**Transparency of safety properties:**

> Since all actions in the alphabet of a property are eligible choices, composing a property with a set of processes *does not affect their correct behaviour*.

> If a behaviour can occur which violates the safety property, then ERROR is reachable.
Transparency

Why must properties be deterministic to be transparent?

Consider:

\[ \text{property } P = (a \rightarrow b \rightarrow P \mid a \rightarrow c \rightarrow P). \]

Is \( a \rightarrow b \) allowed or not?
How can we specify that some action, disaster, never occurs?

A safety property must be specified so as to include all the acceptable, valid behaviours in its alphabet.

property CALM = STOP + \{disaster\}.
Roadmap

- Modelling interaction in FSP
- Safety — synchronizing critical sections
- Locking for atomicity
- The busy-wait mutual exclusion protocol
- Checking Safety properties
- Conditional synchronization
A lock *delays* an acquire request if it is already locked:

\[
\text{LOCK} = (\text{acquire} \rightarrow \text{release} \rightarrow \text{LOCK}).
\]

Similarly, a one-slot buffer delays a put request if it is full and delays a get request if it is empty:

\[
\text{const } N = 2
\]
\[
\text{Slot} = (\text{put}[v:0..N] \\
\rightarrow \text{get}[v] \\
\rightarrow \text{Slot}).
\]
Producer/Consumer composition


Consumer = ( get[x:0..N] -> Consumer ).

||Chain = ( Producer || Slot || Consumer )
Wait and notify

A Java object whose methods are all synchronized behaves like a monitor

Within a synchronized method or block:
> `wait()` suspends the current thread, *releasing the lock*
> `notify()` wakes up *one thread waiting on that object*
> `notifyAll()` *wakes up all threads* waiting on that object

**NB:** Outside of a synchronized block, `wait()` and `notify()` will raise an `IllegalMonitorStateException`

Always use `notifyAll()` unless you are *sure it doesn’t matter which thread you wake up!*
class Slot<
  private Value slotVal; // initially null

public synchronized void put(Value val) {
  while (slotVal != null) {
    try { wait(); } // become NotRunnable
    catch (InterruptedException e) { } } 
  slotVal = val;
  notifyAll(); // make waiting threads Runnable
return;
}

...
public synchronized Value get() {
    Value rval;
    while (slotVal == null) {
        try {
            wait();
        } catch (InterruptedException e) { }
    }
    rval = slotVal;
    slotVal = null;
    notifyAll();
    return rval;
}
Active objects have their own thread.

Producers and Consumers are active objects that communicate and synchronize through a shared buffer.
Active objects

```java
abstract class ActiveObject extends Thread {
    protected int count;
    ActiveObject(String name, int count) {
        super(name);
        this.count = count;
    }
    public void run() {
        int i;
        for (i=1;i<=count;i++) {
            this.action(i);
        }
    }
    protected abstract void action(int n);
}
```

An active object has a thread of its own.
A generic Producer puts count messages to the slot:

abstract class Producer<Value> extends ActiveObject {
    protected Buffer<Value> slot;
    Producer(String name, int count, Buffer<Value> slot) {
        super(name, count);
        this.slot = slot;
    }
    protected void action(int n) {
        slot.put(produce(n));
    }
    protected abstract Value produce(int n);
}
abstract class Consumer<Value> extends ActiveObject {
    protected Buffer<Value> slot;
    Consumer(String name, int count, Buffer<Value> slot) {
        super(name, count);
        this.slot = slot;
    }
    protected void action(int n) {
        consume(n, slot.get());
    }
    protected abstract void consume(int n, Value val);
}
public class ProducerConsumerDemo {
    ...
    private class FruitProducer extends Producer<String> {
        protected String wares;
        FruitProducer(...) { ... }
        protected String produce(int n) {
            String message;
            message = wares + "(" + n + ")";
            System.out.println(getName() + " put " + message);
            return message;
        }
    }
    private class FruitConsumer extends Producer<String> {
        ...
    }
}
Composing Producers and Consumers

Multiple producers and consumers may share the buffer:

```java
public class ProducerConsumerDemo {
    static int COUNT = 5;
    ...
    public void demo() {
        Buffer<String> slot = new Slot<String>();

        new FruitProducer("Peter", COUNT, slot, "apple").start();
        new FruitProducer("Paula", COUNT, slot, "orange").start();
        new FruitProducer("Patricia", COUNT, slot, "banana").start();

        new FruitConsumer("Carla", COUNT, slot).start();
        new FruitConsumer("Cris", 2*COUNT, slot).start();
    }
    ...
}
```

Peter put apple (1)
Carla got apple (1)
Paula put orange(1)
Cris got orange(1)
Patricia put banana(1)
Carla got banana(1)
Peter put apple (2)
Cris got apple (2)
Patricia put banana(2)
Carla got banana(2)
Peter put apple (3)
Cris got apple (3)
Paula put orange(2)
Cris got orange(2)
Patricia put banana(3)
Cris got banana(3)
Peter put apple (4)
Cris got apple (4)
Peter put apple (5)
Cris got apple (5)
Paula put orange(3)
Cris got orange(3)
Patricia put banana(4)
Cris got banana(4)
Patricia put banana(5)
Cris got banana(5)
Paula put orange(4)
Cris got orange(4)
Paula put orange(5)
Cris got orange(5)
What you should know!

> How do you model interaction with FSP?
> What is a critical section? What is critical about it?
> Why don’t sequential programs need synchronization?
> How do locks address safety problems?
> What primitives do you need to implement the busy-wait mutex protocol?
> How can you use FSP to check for safety violations?
> What happens if you call wait or notify outside a synchronized method or block?
> When is it safe to use notifyAll()?
> What are safety properties? How are they modelled in FSP?
Can you answer these questions?

> What is an example of an invariant that might be violated by interfering, concurrent threads?
> What constitute atomic actions in Java?
> Can you ensure safety in concurrent programs without using locks?
> When should you use synchronize(this) rather than synchronize(someObject)?
> Is the busy-wait mutex protocol fair? Deadlock-free?
> How would you implement a Lock class in Java?
> Why is the Java Slot class so much more complex than the FSP Slot specification?
> How would you manually check a safety property?
> Why must safety properties be deterministic to be transparent?
Attribution-ShareAlike 3.0

You are free:
- to copy, distribute, display, and perform the work
- to make derivative works
- to make commercial use of the work

Under the following conditions:

**BY:** Attribution. You must attribute the work in the manner specified by the author or licensor.

**Share Alike.** If you alter, transform, or build upon this work, you may distribute the resulting work only under a license identical to this one.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

http://creativecommons.org/licenses/by-sa/3.0/