Software Engineering using Formal Methods Modeling Concurrency

Wolfgang Ahrendt

08 September 2015

Concurrency: different processes trying not to run into each others' way

Concurrency: different processes trying not to run into each others' way

Main problem of concurrency: sharing computational resources

http://www.youtube.com/watch?v=JgMB6nEv7K0 http://www.youtube.com/watch?v=G8eqymwUFi8

Concurrency: different processes trying not to run into each others' way

Main problem of concurrency: sharing computational resources

http://www.youtube.com/watch?v=JgMB6nEv7K0 http://www.youtube.com/watch?v=G8eqymwUFi8

Shared resource = crossing, bikers = processes,

Concurrency: different processes trying not to run into each others' way

Main problem of concurrency: sharing computational resources

```
http://www.youtube.com/watch?v=JgMB6nEv7K0
http://www.youtube.com/watch?v=G8eqymwUFi8
```

Shared resource = crossing, bikers = processes, and a (data) race in progress, approaching a disaster.

Concurrency: different processes trying not to run into each others' way

Main problem of concurrency: sharing computational resources

```
http://www.youtube.com/watch?v=JgMB6nEv7K0
http://www.youtube.com/watch?v=G8eqymwUFi8
```

Shared resource = crossing, bikers = processes, and a (data) race in progress, approaching a disaster.

Solutions to this must be carefully designed and verified, otherwise. . .



Aim of $\operatorname{SPIN}\text{-style}$ model checking methodology:

exhibit

flaws in

software systems

Aim of Spin-style model checking methodology:

exhibit design flaws in

software systems

Aim of Spin-style model checking methodology:

exhibit design flaws in concurrent and distributed software systems

Aim of Spin-style model checking methodology:

exhibit design flaws in concurrent and distributed software systems

Focus of this lecture:

Modeling and analyzing concurrent systems

Aim of Spin-style model checking methodology:

exhibit design flaws in concurrent and distributed software systems

Focus of this lecture:

Modeling and analyzing concurrent systems

Focus of next lecture:

Modeling and analyzing distributed systems

problems:

▶ hard to predict, hard to form faithful intuition

- hard to predict, hard to form faithful intuition
- enormous combinatorial explosion of possible behavior

- hard to predict, hard to form faithful intuition
- enormous combinatorial explosion of possible behavior
- interleaving prone to unsafe operations

- hard to predict, hard to form faithful intuition
- enormous combinatorial explosion of possible behavior
- interleaving prone to unsafe operations
- counter measures prone to deadlocks

- hard to predict, hard to form faithful intuition
- enormous combinatorial explosion of possible behavior
- ▶ interleaving prone to unsafe operations
- counter measures prone to deadlocks
- ▶ limited control—from within applications—over 'external' factors:

- hard to predict, hard to form faithful intuition
- enormous combinatorial explosion of possible behavior
- interleaving prone to unsafe operations
- counter measures prone to deadlocks
- ▶ limited control—from within applications—over 'external' factors:
 - scheduling strategies

- hard to predict, hard to form faithful intuition
- enormous combinatorial explosion of possible behavior
- interleaving prone to unsafe operations
- counter measures prone to deadlocks
- ▶ limited control—from within applications—over 'external' factors:
 - scheduling strategies
 - relative speed of components

- hard to predict, hard to form faithful intuition
- enormous combinatorial explosion of possible behavior
- ▶ interleaving prone to unsafe operations
- counter measures prone to deadlocks
- ▶ limited control—from within applications—over 'external' factors:
 - scheduling strategies
 - relative speed of components
 - performance of communication mediums

- hard to predict, hard to form faithful intuition
- enormous combinatorial explosion of possible behavior
- interleaving prone to unsafe operations
- counter measures prone to deadlocks
- ▶ limited control—from within applications—over 'external' factors:
 - scheduling strategies
 - relative speed of components
 - performance of communication mediums
 - reliability of communication mediums

Testing Concurrent or Distributed System is Hard

We cannot exhaustively test concurrent/distributed systems

- ► lack of controllability
 - \Rightarrow we miss failures in test phase

Testing Concurrent or Distributed System is Hard

We cannot exhaustively test concurrent/distributed systems

- ► lack of controllability
 - \Rightarrow we miss failures in test phase
- lack of reproducability
 - ⇒ even if failures appear in test phase, often impossible to analyze/debug defect

Testing Concurrent or Distributed System is Hard

We cannot exhaustively test concurrent/distributed systems

- ▶ lack of controllability
 - \Rightarrow we miss failures in test phase
- lack of reproducability
 - ⇒ even if failures appear in test phase, often impossible to analyze/debug defect
- ► lack of time exhaustive testing exhausts the testers long before it exhausts behavior of the system...

Mission of Spin-style Model Checking

offer an efficient methodology to

- ▶ improve the design
- exhibit defects

of concurrent and distributed systems

Activities in Spin-style Model Checking

- 1. model (critical aspects of) concurrent/distributed system with PROMELA
- 2. state crucial properties with assertions, temporal logic, ...
- 3. use SPIN to check all possible runs of the model
- 4. analyze result, possibly re-work 1. and 2.

Activities in Spin-style Model Checking

- 1. model (critical aspects of) concurrent/distributed system with PROMELA
- 2. state crucial properties with assertions, temporal logic, . . .
- 3. use SPIN to check all possible runs of the model
- 4. analyze result, possibly re-work 1. and 2.

Seprate concerns of model vs. property! Check the property you want the model to have, not the one it happens to have.

Main Challenges of Modeling

expressiveness

model must be expressive enough to 'embrace' defects the real system could have

simplicity

model must be simple enough to be 'model checkable', theoretically and practically

Modeling Concurrent Systems in Promela

in the Spin approach, the cornerstone of modeling concurrent/distributed systems are

Promela processes

Initializing Processes

there is always an initial process prior to all others often declared *implicitly* using 'active'

Initializing Processes

there is always an initial process prior to all others often declared *implicitly* using 'active'

```
can be declared explicitly with key word 'init'
init {
   printf("Hello_world\n")
}
```

if explicit, init is used to start other processes with run statement

Starting Processes

```
processes can be started explicitly using run
proctype P() {
  byte local;
init {
  run P();
  run P()
}
each run operator starts copy of process (with copy of local variables)
```

Starting Processes

```
processes can be started explicitly using run
proctype P() {
  byte local;
init {
  run P();
  run P()
}
each run operator starts copy of process (with copy of local variables)
run P() does not wait for P to finish
```

Starting Processes

```
processes can be started explicitly using run
proctype P() {
  byte local;
init {
  run P();
  run P()
}
each run operator starts copy of process (with copy of local variables)
run P() does not wait for P to finish
```

(Promela's run corresponds to Java's start, not to Java's run)

Atomic Start of Multiple Processes

by convention, run operators enclosed in atomic block

```
proctype P() {
   byte local;
   ...
}
init {
   atomic {
    run P();
    run P()
   }
}
```

Atomic Start of Multiple Processes

by convention, run operators enclosed in atomic block

```
proctype P() {
   byte local;
   ...
}
init {
   atomic {
    run P();
    run P()
   }
}
```

effect: processes only start executing once all are created

Atomic Start of Multiple Processes

by convention, run operators enclosed in atomic block

```
proctype P() {
   byte local;
   ...
}
init {
   atomic {
    run P();
    run P()
   }
}
```

effect: processes only start executing once all are created

(more on atomic later)

Joining Processes

```
following trick allows 'joining', i.e., waiting for all processes to finish
byte result;
proctype P() {
init {
  atomic {
    run P();
    run P()
  (_nr_pr == 1); /*blocks until join*/
  printf("result<sub>□</sub>=%d", result)
```

Joining Processes

```
following trick allows 'joining', i.e., waiting for all processes to finish
byte result;
proctype P() {
init {
  atomic {
     run P();
     run P()
  (_nr_pr == 1); /*blocks until join*/
  printf("result<sub>||</sub>=%d", result)
          built-in variable holding number of running processes
_nr_pr
_nr_pr == 1 only 'this' process (init) is (still) running
```

 SEFM: Spin
 CHALMERS/GU
 150908
 14 / 43

Process Parameters

Processes may have formal parameters, instantiated by run:

```
proctype P(byte id; byte incr) {
    ...
}

init {
    run P(7, 10);
    run P(8, 15)
}
```

Active (Sets of) Processes

init can be made implicit by using the active modifier:

```
active proctype P() {
    ...
}
implicit init will run one copy of P
```

Active (Sets of) Processes

init can be made implicit by using the active modifier:

```
active proctype P() {
    ...
}
implicit init will run one copy of P

active [n] proctype P() {
    ...
}
implicit init will run n copies of P
```

Local and Global Data

Variables declared outside of the processes are global to all processes.

Variables declared inside a process are local to that processes.

```
byte n;
proctype P(byte id; byte incr) {
  byte t;
  ...
}
n is global
t is local
```

Modeling with Global Data

pragmatics of modeling with global data:

- **shared memory** of concurrent systems often modeled by global variables of numeric (or array) type
- status of shared resources (printer, traffic light, ...) often modeled by global variables of Boolean or enumeration type (bool/mtype).
- **communication mediums** of distributed systems often modeled by global variables of channel type (chan). (next lecture)

Interference on Global Data

```
byte n = 0;
active proctype P() {
  n = 1;
  printf("Process_P,_n_=_%d\n", n)
}
```

```
byte n = 0;
active proctype P() {
  n = 1;
  printf("Process_P, \_n \_n \_ \_n', n)
}
active proctype Q() {
  n = 2;
  printf("Process_Q, \_n \_n \_ \_n', n)
}
```

```
byte n = 0;
active proctype P() {
  n = 1;
  printf("Process_P, _n_= %d\n", n)
active proctype Q() {
  n = 2;
  printf("Process_{\square}Q,_{\square}n_{\square}=_{\square}%d\n", n)
how many outputs possible?
```

SEFM: SPIN CHALMERS/GU 150908 19 / 43

Interference on Global Data

```
byte n = 0;
active proctype P() {
   n = 1;
   printf("Process_{\square}P,_{\square}n_{\square}=_{\square}%d\n", n)
active proctype Q() {
  n = 2;
   printf("Process_{\square}Q,_{\square}n_{\square}=_{\square}%d\n", n)
```

how many outputs possible?

different processes can interfere on global data

Examples

- interleave0.pml
 SPIN simulation, SPINSPIDER automata + transition system
- interleave1.pml
 SPIN simulation, adding assertion, fine-grained execution model,
 model checking
- 3. interleave5.pml SPIN simulation, SPIN model checking, trail inspection

Atomicity

limit the possibility of sequences being interrupted by other processes

weakly atomic sequence

can only be interrupted if a statement is not executable

strongly atomic sequence

cannot be interrupted at all

Atomicity

```
limit the possibility of sequences being interrupted by other processes

weakly atomic sequence

can only be interrupted if a statement is not executable

defined in PROMELA by atomic{ ... }

strongly atomic sequence

cannot be interrupted at all

defined in PROMELA by d_step{ ... }
```

Deterministic Sequences

$d_step:$

- strongly atomic
- deterministic (like a single step)

Deterministic Sequences

$d_step:$

- strongly atomic
- deterministic (like a single step)
- choices resolved in fixed way (always take the first possible option)
 - \Rightarrow avoid choices in d_step

Deterministic Sequences

d_{step} :

- strongly atomic
- deterministic (like a single step)
- choices resolved in fixed way (always take the first possible option)
 ⇒ avoid choices in d_step
- it is an error if any statement within d_step, other than the first one (called 'guard'), blocks

```
\begin{array}{ll} \mathbf{d\_step} & \{ \\ & \mathtt{stmt1}; \; \leftarrow \; \mathit{guard} \\ & \mathtt{stmt2}; \\ & \mathtt{stmt3} \\ \} \end{array}
```

If stmt1 blocks, d_step is not entered, and blocks as a whole.

It is an error if stmt2 or stmt3 block.

(Weakly) Atomic Sequences

atomic:

- weakly atomic
- can be non-deterministic

```
atomic {
    stmt1; ← guard
    stmt2;
    stmt3
}
```

If guard blocks, atomic is not entered, and blocks as a whole.

Once atomic is entered, control is kept until a statement blocks, and only then passed to another process.

Prohibit Interference by Atomicity

apply atomic or d_step to interference examples

PROMELA has no synchronization primitives, like semaphores, locks, or monitors.

PROMELA has no synchronization primitives, like semaphores, locks, or monitors.

Instead, Promela inhibits concept of statement executability

PROMELA has no synchronization primitives, like semaphores, locks, or monitors.

Instead, Promela inhibits concept of statement executability

Executability addresses many issues in the interplay of processes

PROMELA has no synchronization primitives, like semaphores, locks, or monitors.

Instead, Promela inhibits concept of statement executability

Executability addresses many issues in the interplay of processes

Most known synchronization primitives (e.g. test & set, compare & swap, semaphores) can be modelled using executability and atomicity

Executability

Each statement has the notion of executability. Executability of basic statements:

statement type	executable
assignment	always
assertion	always
print statement	always
expression statement	iff value not $0/false$
send/receive statement	(next lecture)

Executability of compound statements:

Executability of compound statements:

 $\begin{array}{c} \mathbf{atomic} \ \mathsf{resp.} \ \mathbf{d_step} \ \mathsf{statement} \ \mathsf{is} \ \mathsf{executable} \\ \mathsf{iff} \\ \mathsf{guard} \ \mathsf{(i.e., the first inner statement)} \ \mathsf{is} \ \mathsf{executable} \end{array}$

Executability of compound statements:

```
atomic resp. d_step statement is executable iff guard (i.e., the first inner statement) is executable
```

 $\begin{array}{c} \textbf{if resp. do statement is executable} \\ & \textbf{iff} \end{array}$

any of its alternatives is executable

Executability of compound statements:

```
atomic resp. d_step statement is executable
guard (i.e., the first inner statement) is executable
        if resp. do statement is executable
                         iff
        any of its alternatives is executable
            an alternative is executable
                         iff
   its guard (the first statement) is executable
```

Executability of compound statements:

```
atomic resp. d_step statement is executable
 guard (i.e., the first inner statement) is executable
         if resp. do statement is executable
                          iff
         any of its alternatives is executable
             an alternative is executable
                          iff
    its guard (the first statement) is executable
(recall: in alternatives, "->" syntactic sugar for ";")
```

Executability and Blocking

Definition (Blocking)

A statement blocks iff it is not executable.

A process blocks iff its location counter points to a blocking statement.

For each step of execution, the scheduler nondeterministically chooses a process to execute

Executability and Blocking

Definition (Blocking)

A statement blocks iff it is not executable.

A process blocks iff its location counter points to a blocking statement.

For each step of execution, the scheduler nondeterministically chooses a process to execute among the non-blocking processes.

Executability and Blocking

Definition (Blocking)

A statement blocks iff it is not executable.

A process blocks iff its location counter points to a blocking statement.

For each step of execution, the scheduler nondeterministically chooses a process to execute among the non-blocking processes.

Executability, resp. blocking are the key to PROMELA-style modeling of solutions to synchronization problems. (to be discussed in the following)

The Critical Section Problem

archetypical problem of concurrent systems given a number of looping processes, each containing a critical section design an algorithm such that:

The Critical Section Problem

archetypical problem of concurrent systems
given a number of looping processes, each containing a critical section

design an algorithm such that:

Mutual Exclusion At most one process is executing its critical section at any time.

The Critical Section Problem

archetypical problem of concurrent systems

given a number of looping processes, each containing a critical section design an algorithm such that:

Mutual Exclusion At most one process is executing its critical section at any time.

Absence of Deadlock If *some* processes are trying to enter their critical sections, then *one* of them must eventually succeed.

The Critical Section Problem

archetypical problem of concurrent systems

given a number of looping processes, each containing a critical section design an algorithm such that:

Mutual Exclusion At most one process is executing its critical section at any time.

Absence of Deadlock If *some* processes are trying to enter their critical sections, then *one* of them must eventually succeed.

Absence of (individual) Starvation If *any* process tries to enter its critical section, then *that* process must eventually succeed.

```
for demonstration and simplicity:
(non)critical sections only printf statements
active proctype P() {
  do :: printf("P<sub>□</sub>non-critical<sub>□</sub>actions\n");
         /* begin critical section */
         printf("Pusesusharedurecourses\n")
         /* end critical section */
  od
active proctype Q() {
  do :: printf("Qunon-criticaluactions\n");
         /* begin critical section */
         printf("Quusesusharedurecourses\n")
         /* end critical section */
  od
```

```
More infrastructure to achieve ME.
Adding two Boolean flags:
bool P_in_CS = false;
bool Q_in_CS = false;
active proctype P() {
  do :: printf("P<sub>||</sub>non-critical<sub>||</sub>actions\n");
          P in CS = true:
          /* begin critical section */
          printf("P<sub>|</sub>uses<sub>|</sub>shared<sub>|</sub>recourses\n");
          /* end critical section */
          P in CS = false
  od
active proctype Q() {
  \dots correspondingly \dots
```

```
adding assertions
bool P_in_CS = false;
bool Q_in_CS = false;
active proctype P() {
  do :: printf("Punon-criticaluactions\n");
        P_{in}CS = true;
        /* begin critical section */
        printf("P_uses_shared_recourses\n");
        assert(!Q_in_CS);
        /* end critical section */
        P_{in}_{CS} = false
  od
active proctype Q() {
    .....assert(!P_in_CS);......
}
```

```
bool P_in_CS = false;
bool Q_in_CS = false;
active proctype P() {
  do :: printf("Punon-criticaluactions\n");
        P_{in}_{CS} = true;
        do :: !Q_in_CS -> break
            :: else -> skip
        od;
        /* begin critical section */
        printf("P_uses_shared_recourses\n");
        assert(!Q_in_CS);
        /* end critical section */
        P_{in}_{CS} = false
  od
active proctype Q() { ...correspondingly... }
```

Mutual Exclusion by Blocking

instead of Busy Waiting, process should

- yield control
- continuing to run only when exclusion properties are fulfilled

Mutual Exclusion by Blocking

instead of Busy Waiting, process should

- yield control
- continuing to run only when exclusion properties are fulfilled

We can use expression statement !Q_in_CS, to let process P block where it should not proceed!

```
active proctype P() {
  do :: printf("Punon-criticaluactions\n");
         P_{in}_{CS} = true;
          !Q_in_CS:
          /* begin critical section */
          printf("P<sub>□</sub>uses<sub>□</sub>shared<sub>□</sub>recourses\n");
          assert(!Q_in_CS);
          /* end critical section */
          P in CS = false
  od
active proctype Q() {
  \dots correspondingly \dots
```

Verify with $\ensuremath{\mathrm{S}}\xspace\mathrm{PIN}$

Verify with SPIN

SPIN error (invalid end state)

 \Rightarrow deadlock

Verify with SPIN

SPIN error (invalid end state) \Rightarrow deadlock

can make pan ignore the deadlock: ./pan -E

```
Verify with SPIN
```

```
SPIN error (invalid end state) \Rightarrow deadlock
```

can make pan ignore the deadlock: ./pan -E

Spin still reports assertion violation(!)

▶ mutual exclusion (ME) cannot be shown by SPIN

- ▶ mutual exclusion (ME) cannot be shown by SPIN
- ► P/Q_in_CS sufficient for achieving ME

- ▶ mutual exclusion (ME) cannot be shown by SPIN
- ▶ P/Q_in_CS sufficient for achieving ME
- ► P/Q_in_CS *not* sufficient for *proving* ME

- ▶ mutual exclusion (ME) cannot be shown by SPIN
- ► P/Q_in_CS sufficient for achieving ME
- ▶ P/Q_in_CS not sufficient for proving ME

need more infrastructure:

ghost variables, only for proving / model checking

Show Mutual Exclusion with Ghost Variable

```
int critical = 0;
active proctype P() {
  do :: printf("Punon-criticaluactions\n");
         P_{in}_{CS} = true;
         ! Q in CS:
         /* begin critical section */
         critical++:
         printf("P_uses_shared_recourses\n");
         assert(critical < 2);</pre>
         critical--:
         /* end critical section */
         P_{in}CS = false
  od
active proctype Q() {
  \dots correspondingly \dots
```

SPIN (./pan -E) shows no assertion is violated ⇒ mutual exclusion is verified

```
SPIN (./pan -E) shows no assertion is violated \Rightarrow mutual exclusion is verified
```

Still SPIN (without -E) reports (invalid end state) \Rightarrow deadlock

Deadlock Hunting

Invalid End State:

- A process does not finish at its end
- OK if it is not crucial to continue see end.pml
- If it is crucial to continue: Real deadlock

Deadlock Hunting

Invalid End State:

- A process does not finish at its end
- ▶ OK if it is not crucial to continue see end.pml
- If it is crucial to continue:
 Real deadlock

Find Deadlock with Spin:

- Verify to produce a failing run trail
- Simulate to see how the processes get to the interlock
- ► Fix the code (not using the end...: labels or -E option)

Atomicity against Deadlocks

solution:

checking and setting the flag in one atomic step

Atomicity against Deadlocks

solution:

checking and setting the flag in one atomic step

(demonstrae that in csGhost.pml)

Atomicity against Deadlocks

solution:

checking and setting the flag in one atomic step

(demonstrae that in csGhost.pml)

```
atomic {
  !Q_in_CS;
  P_in_CS = true
}
```

designated artifacts for verification:

- designated artifacts for verification:
 - ghost variables ('verification only' variables)

- designated artifacts for verification:
 - ghost variables ('verification only' variables)
 - temporal logic (later in the course)

- designated artifacts for verification:
 - ghost variables ('verification only' variables)
 - temporal logic (later in the course)
- max n processes allowed in critical section modeling possibilities include:

- designated artifacts for verification:
 - ghost variables ('verification only' variables)
 - temporal logic (later in the course)
- max n processes allowed in critical section modeling possibilities include:
 - counters instead of booleans

- designated artifacts for verification:
 - ghost variables ('verification only' variables)
 - temporal logic (later in the course)
- max n processes allowed in critical section modeling possibilities include:
 - counters instead of booleans
 - semaphores (see demo)

- designated artifacts for verification:
 - ghost variables ('verification only' variables)
 - temporal logic (later in the course)
- max n processes allowed in critical section modeling possibilities include:
 - counters instead of booleans
 - semaphores (see demo)
- more fine grained exclusion conditions, e.g.

- designated artifacts for verification:
 - ghost variables ('verification only' variables)
 - temporal logic (later in the course)
- max n processes allowed in critical section modeling possibilities include:
 - counters instead of booleans
 - semaphores (see demo)
- more fine grained exclusion conditions, e.g.
 - several critical sections (Leidestraat in Amsterdam)

- designated artifacts for verification:
 - ghost variables ('verification only' variables)
 - temporal logic (later in the course)
- max n processes allowed in critical section modeling possibilities include:
 - counters instead of booleans
 - semaphores (see demo)
- more fine grained exclusion conditions, e.g.
 - several critical sections (Leidestraat in Amsterdam)
 - writers exclude each other and readers readers exclude writers, but not other readers

- designated artifacts for verification:
 - ghost variables ('verification only' variables)
 - temporal logic (later in the course)
- max n processes allowed in critical section modeling possibilities include:
 - counters instead of booleans
 - semaphores (see demo)
- more fine grained exclusion conditions, e.g.
 - several critical sections (Leidestraat in Amsterdam)
 - writers exclude each other and readers readers exclude writers, but not other readers
 - ▶ FIFO queue semaphores, for fairly choosing processes to enter

- designated artifacts for verification:
 - ghost variables ('verification only' variables)
 - temporal logic (later in the course)
- max n processes allowed in critical section modeling possibilities include:
 - counters instead of booleans
 - semaphores (see demo)
- more fine grained exclusion conditions, e.g.
 - several critical sections (Leidestraat in Amsterdam)
 - writers exclude each other and readers readers exclude writers, but not other readers
 - ▶ FIFO queue semaphores, for fairly choosing processes to enter
- ... and many more

Why Not Critical Section in Single Atomic Block?

Why Not Critical Section in Single Atomic Block?

Actually possible in this case.

Also in interleaving example (counting via temp, see above).

But:

- does not carry over to variations (see previous slide)
- ► atomic only weakly atomic!
- ▶ d_step excludes any nondeterminism!

Why Not Critical Section in Single Atomic Block?

Actually possible in this case.

Also in interleaving example (counting via temp, see above).

But:

- does not carry over to variations (see previous slide)
- ► atomic only weakly atomic!
- ▶ d_step excludes any nondeterminism!

Using atomic and d_step too heavily, for too large blocks, can result in well-behaved models, while modelling the wrong system.