Model-Checking Concurrent Systems

Wolfgang Schreiner
Wolfgang.Schreiner@risc.jku.at
Research Institute for Symbolic Computation (RISC)
Johannes Kepler University, Linz, Austria
http://www.risc.jku.at

1. The Model Checker Spin

2. Automatic Model Checking

The Model Checker Spin

- Spin system:
  - Gerard J. Holzmann et al, Bell Labs, 1980–.
  - Freely available since 1991.
  - Symposium series since 1995 (29th symposium “Spin 2023”).
- Spin resources:
- Goal: verification of (concurrent/distributed) software models.

On-the-fly LTL model checking of finite state systems.

- System S modeled by automaton $S_A$.
  - Explicit representation of automaton states.
  - There exist various other approaches (discussed later).
- On-the-fly model checking.
  - Reachable states of $S_A$ are only expended on demand.
  - Partial order reduction to keep state space manageable.
- LTL model checking.
  - Property $P$ to be checked described in PLTL.
    - Propositional linear temporal logic.
  - Description converted into property automaton $P_A$.
    - Automaton accepts only system runs that do not satisfy the property.
- Model checking based on automata theory.
The Spin System Architecture

- System description in Promela.
  - Promela = Process Meta-Language.
  - Spin = Simple Promela Interpreter.
- Express coordination and synchronization aspects of a real system.
- Actual computation can be e.g. handled by embedded C code.

Simulation mode.
- Investigate individual system behaviors.
- Inspect system state.
- Graphical interface ispin for visualization.

Verification mode.
- Verify properties shared by all possible system behaviors.
- Properties specified in PLTL and translated to “never claims”.
- Promela description of automaton for negation of the property.
- Generated counter examples may be investigated in simulation mode.

Features of Spin

Verification and simulation are tightly integrated in Spin.

The Spin User Interface

The Client/Server System in Promela

/ * definition of a constant MESSAGE */
  mtype = { MESSAGE };

/ * two arrays of channels of size 2, 
   each channel has a buffer size 1 */
  chan request[2] = [1] of { mtype };
  chan answer[2] = [1] of { mtype };

/ * the system of three processes */
init
  { run client(1);
    run client(2);
    run server();
  }

/ * the mutual exclusion property */
ltl mutex { []!(client[1]@C && client[2]@C) }
The Client/Server System in Promela

```promela
/* the server process type */
proctype server()
{
    /* three variables of two bit each */
    unsigned given : 2 = 0;
    unsigned waiting : 2 = 0;
    unsigned sender : 2;

do :: true ->
    /* receiving the message */
    R: if
        :: request[0] ? MESSAGE ->
            S1: sender = 1
        :: request[1] ? MESSAGE ->
            S2: sender = 2
    fi;

    /* answering the message */
    if
        :: sender == given ->
            if
                :: waiting == 0 ->
                    given = 0
                :: else ->
                    given = waiting;
                    waiting = 0;
                    answer[given-1] ! MESSAGE
            fi;
        :: given == 0 ->
            given = sender;
            answer[given-1] ! MESSAGE
        :: else
            waiting = sender
    fi;

    od;
}
```

Simulating the System Execution in Spin

The Spin Simulation View

The Spin Verification View
Specifying a System Property in Spin

The easiest way to specify an LTL property is to specify it inline. The formula is specified globally (i.e., outside all proctype or init declarations) with the following syntax:

\[ \text{ltl } \{ \text{name} \} \{ \{ \text{formula} \} \} \]

... The names of operators can either be abbreviated with the symbols shown above, or spelled out in full (as always, eventually, until, implies, and equivalent). The alternative operators weakuntil, stronguntil, and release (for the V operator, see above), are also supported. This means that the following two are equivalent:

\[ \text{ltl } p1 \{ []<> p \} \]
\[ \text{ltl } p2 \{ \text{always eventually } p \} \]
\[ \text{ltl } p3 \{ \text{eventually } (a > b) \text{ implies } \text{len}(q) == 0 \} \]

There is just one restriction: you cannot use the predefined operators empty, nempty, full, nfull in inline ltl formula...

Spin LTL

Grammar:

\[ \text{ltl} ::= \text{opd} | ( \text{ltl} ) | \text{ltl } \text{binop} \text{ ltl} | \text{unop} \text{ ltl} \]

Operands (opd):

- true, false, user-defined names starting with a lower-case letter, or embedded expressions inside curly braces, e.g.,: \{ a+b>n \}.

Unary Operators (unop):

- [] (the temporal operator always)
- <> (the temporal operator eventually)
- ! (the boolean operator for negation)

Binary Operators (binop):

- U (the temporal operator strong until)
- W (the temporal operator weak until)
- V (the dual of U): \( (p \text{ V } q) \text{ means } !(!p \text{ U } !q) \)
- && (the boolean operator for logical and)
- || (the boolean operator for logical or)
- /\ (alternative form of &&)
- /\ (alternative form of ||)
- -> (the boolean operator for logical implication)
- <= (the boolean operator for logical equivalence)

Spin Atomic Predicates

(a+b > c)
(len(q) < 5)
(process@Label)
(process[pid]@Label)

- PROMELA conditions with references to global system variables.

  C-like boolean expressions.

  - len(q): the number of messages in channel q.
  - process@Label: true if the execution of the process with process type process is in the state marked by Label.
  - process[pid]@Label: true if the execution of the process with type process and process identifier pid is in the state marked by Label.

  - Process init receives identifier 0.
  - Processes instantiated with run receive identifiers 1,2,...

Atomic predicates can describe arbitrary state conditions.

The Spin Verification Options

- Verification mode
  - Safety: predefined safety properties.
  - Liveness: predefined liveness properties.
  - Use Claim: user-defined (LTL) properties.

- Storage mode
  - Exhaustive: full error search.
    - May require (too) much memory.
  - Hash-compact and bitstate/supertrace: probabilistic search.
    - Less memory consuming but not guaranteed to find all errors.
    - Number of reported hash collisions should not be too high.

- Advanced Parameters
  - Physical Memory Size: increase, if checker runs out of memory.
  - Maximum Search Depth: reduce to find shorter counterexamples.

By default, select “use claim” and “exhaustive”.

Wolfgang Schreiner 13/47
http://www.risc.jku.at

Wolfgang Schreiner 14/47
http://www.risc.jku.at

Wolfgang Schreiner 15/47
http://www.risc.jku.at

Wolfgang Schreiner 16/47
http://www.risc.jku.at
The Spin Verification Output

(Spin Version 6.1.0 -- 4 May 2011)
+ Partial Order Reduction

Full statespace search for:
never claim + (mutex)
assertion violations + (if within scope of claim)
acceptance cycles + (fairness disabled)
invalid end states - (disabled by never claim)

State-vector 48 byte, depth reached 137, errors: 0
157 states, stored
68 states, matched
225 transitions (= stored+matched)
0 atomic steps
hash conflicts: 0 (resolved)

Stats on memory usage (in Megabytes):
...
pan: elapsed time 0 seconds
No errors found -- did you verify all claims?

Wolfgang Schreiner http://www.risc.jku.at 17/47

More Promela Features

Active processes, inline definitions, atomic statements, output.
mtype = { P, C, N }
mtype turn = P;
inline request(x, y) { atomic { x == y -> x = N } }
inline release(x, y) { atomic { x = y } }
#define FORMAT "Output: %c\n"

active proctype producer()
{
    do
    :: request(turn, P) -> printf(FORMAT, 'P'); release(turn, C);
    od
}

active proctype consumer()
{
    do
    :: request(turn, C) -> printf(FORMAT, 'C'); release(turn, P);
    od
}

Wolfgang Schreiner http://www.risc.jku.at 18/47

1. The Model Checker Spin

2. Automatic Model Checking

Can embed computational aspects into a Promela model (only works in verification mode where a C program is generated from the model).

Wolfgang Schreiner http://www.risc.jku.at 19/47
The Basic Approach

Translation of the original problem to a problem in automata theory.

- Original problem: \( S \models P \).
  - \( S = \langle I, R \rangle \), PLTL formula \( P \).
  - Does property \( P \) hold for every run of system \( S \)?
- Construct system automaton \( S_A \) with language \( L(S_A) \).
  - A language is a set of infinite words.
  - Each such word describes a system run.
  - \( L(S_A) \) describes the set of runs of \( S \).
- Construct property automaton \( P_A \) with language \( L(P_A) \).
  - \( L(P_A) \) describes the set of runs satisfying \( P \).
- Equivalent Problem: \( L(S_A) \subseteq L(P_A) \).
  - The language of \( S_A \) must be contained in the language of \( P_A \).

There exists an efficient algorithm to solve this problem.

Finite State Automata

A (variant of a) labeled transition system in a finite state space.

- Take finite sets \( State \) and \( Label \).
  - The state space \( State \).
  - The alphabet \( Label \).
- A (finite state) automaton \( A = \langle I, R, F \rangle \) over \( State \) and \( Label \):
  - A set of initial states \( I \subseteq State \).
  - A labeled transition relation \( R \subseteq Label \times State \times State \).
  - A set of final states \( F \subseteq State \).
    - Büchi automata: \( F \) is called the set of accepting states.

We will only consider infinite runs of Büchi automata.

A Finite State System as an Automaton

The automaton \( S_A = \langle I, R, F \rangle \) for a finite state system \( S = \langle I_S, R_S \rangle \):

- \( State := State_S \cup \{ \iota \} \).
  - The state space \( State_S \) of \( S \) is finite; additional state \( \iota \) ("iota").
- \( Label := \mathbb{P}(AP) \).
  - Finite set \( AP \) of atomic propositions.
    - All PLTL formulas are built from this set only.
  - Power set \( \mathbb{P}(S) := \{ s : s \subseteq S \} \).
  - Every element of \( Label \) is thus a set of atomic propositions.
- \( I := \{ \iota \} \).
  - Single initial state \( \iota \).
- \( R(I, s, s') :\Leftrightarrow I = L(s') \land (R_S(s, s') \lor (s = \iota \land I_S(s'))) \).
  - \( L(s) := \{ p \in AP : s \models p \} \).
  - Each transition is labeled by the set of atomic propositions satisfied by the successor state.
- \( F := State \).
  - Every state is accepting.
A Finite State System as an Automaton

If \( r = s_0 \rightarrow s_1 \rightarrow s_2 \rightarrow \ldots \) is a run of \( S \), then \( S_A \) accepts the labelled version \( r_l := \iota \left( L(s_0) \rightarrow s_0 \right) \left( L(s_1) \rightarrow s_1 \right) \left( L(s_2) \rightarrow s_2 \right) \ldots \) of \( r \).

Further Examples

- Example: \( \Diamond \square p \).
- Example: \( \square \Diamond p \).

Arbitrary PLTL formulas can be converted to automata (we omit the details).

System Properties

- State equivalence: \( L(s) = L(t) \).
- Both states have the same labels.
- Both states satisfy the same atomic propositions in \( AP \).
- Run equivalence: \( w(r_l) = w(r'_l) \).
- Both runs have the same sequences of labels.
- Both runs satisfy the same PLTL formulas built over \( AP \).
- Indistinguishability: \( w(r_l) = w(r'_l) \Rightarrow (r \models P \Leftrightarrow r' \models P) \)
- PLTL formula \( P \) cannot distinguish between runs \( r \) and \( r' \) whose labeled versions read the same words.
- Consequence: \( S \models P \Leftrightarrow \mathcal{L}(S_A) \subseteq \mathcal{L}(P_A) \).
- If every run of \( S \) satisfies \( P \), then every word \( w(r_l) \) in \( \mathcal{L}(S_A) \) equals some word \( w(r'_l) \) in \( \mathcal{L}(P_A) \), and vice versa.
The Next Steps

- Problem: \( L(S_A) \subseteq L(P_A) \)
  - Equivalent to: \( L(S_A) \cap L(P_A) = \emptyset \).
  - Complement \( L := \{w : w \not\in L\} \).
  - Equivalent to: \( L(S_A) \cap L(\neg P_A) = \emptyset \).
- Equivalent Problem: \( L(S_A) \cap L(\neg P_A) = \emptyset \).

We will introduce the synchronized product automaton \( A \otimes B \).

A transition of \( A \otimes B \) represents a simultaneous transition of \( A \) and \( B \).

- Property: \( L(A) \cap L(B) = L(A \otimes B) \).
- Final Problem: \( L(S_A \otimes (\neg P_A)) = \emptyset \).

We have to check whether the language of this automaton is empty.
We have to look for a word \( w \) accepted by this automaton.
- If no such \( w \) exists, then \( S \models P \).
- If such a \( w = w(r) \) exists, then \( r \) is a counterexample, i.e. a run of \( S \) such that \( r \not\models P \).

Synchronized Product of Two Automata

Given two finite automata \( A = \langle I_A, R_A, State_A \rangle \) and \( B = \langle I_B, R_B, F_B \rangle \).

- Synchronized product \( A \otimes B = \langle I, R, F \rangle \).
  - State := \( State_A \times State_B \).
  - Label := \( Label_A = Label_B \).
  - \( I := I_A \times I_B \).
  - \( R(l, (s_A, s_B), (s'_A, s'_B)) :\iff R_A(l, s_A, s'_A) \land R_B(l, s_B, s'_B) \).
  - \( F := State_A \times F_B \).

Special case where all states of automaton \( A \) are accepting.

Example

Check whether \( S \models \Box(P \Rightarrow \Diamond Q) \).


The product automaton accepts a run, thus the property does not hold (shown \( S \) has no \( \iota \) but \( t_i \) labeled with properties of predecessor state).
Checking Emptiness

How to check whether \( L(A) \) is non-empty?

- If \( L(A) \) is non-empty, \( A \) accepts some run \( r \).
  - \( r \) represents a counterexample for the property to be checked.
- Since \( r \) is accepted, it contains infinitely many occurrences of some accepting state \( s \).
  - Finite prefix \( l \to \ldots s \to \ldots s \to \ldots \)
  - Since the state space is finite, \( r \) must contain a cycle \( s \to \ldots s \).
  - Infinite repetition of cycle \( s \to \ldots s \).

We have to search for such an acceptance cycle.

An accepting state \( s \) that is reachable from itself.

The search for an acceptance cycle in the reachability graph is the core problem of PLTL model checking; it can be solved by depth-first search.

Searching for a State

Algorithm to check violation of assertion \( f \) (i.e., \( \square f \) with state formula \( f \)).

```plaintext
global
StateSpace V := {}
Stack D := ()
proc main()
  // r becomes true, iff
  // counterexample run is found
  push(D, i)
  r := search(s)
  pop(D)
end

function search(s)
  V := V \cup \{ s \}
  if s \neq f then
    print D
    return true
  end
  for (l, s, s') \in R do
    if s' \notin V
      push(D, s')
      r := search(s')
      pop(D)
    if r then return true end
  end
  return false
end
```

Stack \( D \) can be used to print counterexample run.

Basic Structure of Depth-First Search

Visit all states of the reachability graph of an automaton \( \{ \{ s \}, R, F \} \).

```plaintext
global
StateSpace V := {}
Stack D := ()
proc visit(s)
  V := V \cup \{ s \}
  for (l, s, s') \in R do
    if s' \notin V
      push(D, s')
      visit(s')
      pop(D)
    end
  end
end
```

State space \( V \) holds all states visited so far; stack \( D \) holds path from initial state to currently visited state.

Searching for an Acceptance Cycle

Algorithm to check whether a general temporal formula is violated.

```plaintext
global
StateSpace V := {}
Stack D := ()
proc main()
  push(D, i); r := search(s); pop(D)
end

function searchCycle(s)
  for (l, s, s') \in R do
    if has(D, s') then
      print D; print C; print s'
      return true
    end
    else if ¬has(C, s') then
      push(C, s');
      r := searchCycle(s')
      pop(C);
      if r then return true end
    end
  end
  return false
end
```
**Complexity of the Search**

The complexity of checking $S \models P$ is as follows.
- Let $|P|$ denote the number of subformulas of $P$.
- $|\text{State}_{\neg P}|_A = O(2^{|P|})$.
- $|\text{State}_{A \otimes B}| = |\text{State}_A| \cdot |\text{State}_B|$.
- $|\text{State}_{S_A \otimes \neg P}|_A = O(|\text{State}_{S_A}| \cdot 2^{|P|})$
- The time complexity of $\text{search}$ is linear in the size of $\text{State}$.
  - Actually, in the number of reachable states (typically much smaller).
  - Only true for the improved variant where the $c$ bits are not reset.
  - Then every state is visited by $\text{searchCycle}$ at most once.

PLTL model checking is linear in the number of reachable states but exponential in the size of the formula.

**The Overall Process**

Basic PLTL model checking for deciding $S \models P$.
- Convert system $S$ to automaton $S_A$.
  - Atomic propositions of PLTL formula are evaluated on each state.
- Convert negation of PLTL formula $P$ to automaton $(\neg P)_A$.
  - How to do so, we have not yet described.
- Construct synchronized product automaton $S_A \otimes (\neg P)_A$.
  - After that, formula labels are not needed any more.
- Find acceptance cycle in reachability-graph of product automaton.
  - A purely graph-theoretical problem that can be efficiently solved.
  - Time complexity is linear in the size of the state space of the system but exponential in the size of the formula to be checked.
  - Weak scheduling fairness with $k$ components: runtime is increased by factor $k + 2$ (worst-case, “in practice just factor 2” [Holzmann]).

The basic approach immediately leads to state space explosion; further improvements are needed to make it practical.

**Implementing the Search**

- The state space $V$, is implemented by a hash table for efficiently checking $s' \notin V$.
- Rather than using explicit stacks $D$ and $C$,
  - each state node has two bits $d$ and $c$,
  - $d$ is set to denote that the state is in stack $D$,
  - $c$ is set to denote that the state is in stack $C$.
- The counterexample is printed.
  - by searching, starting with $\iota$, the unique sequence of reachable nodes where $d$ is set until the accepting node $s_a$ is found, and
  - by searching, starting with a successor of $s_a$, the unique sequence of reachable nodes where $c$ is set until the cycle is detected.
- Furthermore, it is not necessary to reset the $c$ bits, because
  - $\text{search}$ first explores all states reachable by an accepting state $s$ before trying to find a cycle from $s$; from this, one can show that
called with first accepting node $s_a$ reachable from itself, $\text{searchCycle}$ needs not revisit nodes with $c$ bits set in previous searches.
  - With this improvement, every state is only visited once.

**Depth-First Search for Acceptance Cycle**

- At each call of $\text{search}(s)$,
  - $s$ is a reachable state,
  - $D$ describes a path from $\iota$ to $s$.
- $\text{search}$ calls $\text{searchCycle}(s)$
  - on a reachable accepting state $s$
  - in order to find a cycle from $s$ to itself.
- At each call of $\text{searchCycle}(s)$,
  - $s$ is a state reachable from a reachable accepting state $s_a$,
  - $D$ describes a path from $\iota$ to $s_a$,
  - $D \otimes C$ describes a path from $\iota$ to $s$ (via $s_a$).
- Thus we have found an accepting cycle $D \otimes C \rightarrow s'$, if
  - there is a transition $s \not\rightarrow s'$,
  - such that $s'$ is contained in $D$.

If the algorithm returns “true”, there exists a violating run; the converse follows from the exhaustiveness of the search.
On the Fly Model Checking

For checking $L(S_A \otimes (\neg P)_A) = \emptyset$, it is not necessary to construct the states of $S_A$ in advance.

- Only the property automaton $(\neg P)_A$ is constructed in advance.
  - This automaton has comparatively small state space.
- The system automaton $S_A$ is constructed on the fly.
  - Construction is guided by $(\neg P)_A$ while computing $S_A \otimes (\neg P)_A$.
  - Only that part of the reachability graph of $S_A$ is expanded that is consistent with $(\neg P)_A$ (i.e. can lead to a counterexample run).
- Typically only a part of the state space of $S_A$ is investigated.
  - A smaller part, if a counterexample run is detected early.
  - A larger part, if no counterexample run is detected.

Unreachable system states and system states that are not along possible counterexample runs are never constructed.

Example

Check $(T_1 \parallel T_2) \models \Diamond g \geq 2$.

For checking $\Diamond g \geq 2$, it suffices to check only one ordering of the independent transitions $x = 1$ and $y = 1$ (not true for checking $\Box x \geq y$).

Partial Order Reduction

Core problem of model checking: state space explosion.

- Take asynchronous composition $S_0 \parallel S_1 \parallel \ldots \parallel S_{k-1}$.
- Take state $s$ where one transition of each component is enabled.
  - Assume that the transition of one component does not disable the transitions of the other components and that no other transition becomes enabled before all the transitions have been performed.
- Take state $s'$ after execution of all the transitions.
  - There are $k!$ paths leading from $s$ to $s'$.
  - There are $2^k$ states involved in the transitions.

Sometimes it suffices to consider a single path with $k + 1$ states.
Other Optimizations

- **Statement merging.**
  - Special case of partial order reduction where a sequence of transitions of same component is combined to a single transition.

- **State compression.**
  - **Collapse compression:** each state holds pointers to component states; thus component states can be shared among many system states.
  - **Minimized automaton representation:** represent state set \( V \) not by hash table but by finite state automaton that accepts a state (sequence of bits) \( s \) if and only if \( s \in V \).
  - **Hash compact:** store in the hash table a hash value of the state (computed by a different hash function). Probabilistic approach: fails if two states are mapped to the same hash value.
  - **Bitstate hashing:** represent \( V \) by a bit table whose size is much larger than the expected number of states; each state is then only represented by a single bit. Probabilistic approach: fails if two states are hashed to the same position in the table.

Other Approaches to Model Checking

There are fundamentally different approaches to model checking than the automata-based one implemented in Spin.

- **Symbolic Model Checking (e.g. SMV, NuSMV).**
  - Core: binary decision diagrams (BDDs).
    - Data structures to represent boolean functions.
    - Can be used to describe state sets and transition relations.
  - The set of states satisfying a CTL formula \( P \) is computed as the BDD representation of a fixpoint of a function (predicate transformer) \( F_P \).
    - If all initial system states are in this set, \( P \) is a system property.
  - BDD packages for efficiently performing the required operations.

- **Bounded Model Checking (e.g. NuSMV2).**
  - Core: propositional satisfiability.
    - Is there a truth assignment that makes propositional formula true?
    - There is a counterexample of length at most \( k \) to a LTL formula \( P \), if and only if a particular propositional formula \( F_{k,P} \) is satisfiable.
    - Problem: find suitable bound \( k \) that makes method complete.
  - SAT solvers for efficiently deciding propositional satisfiability.

- **Counter-Example Guided Abstraction Refinement (e.g. BLAST).**
  - Core: model abstraction.
    - A finite set of predicates is chosen and an abstract model of the system is constructed as a finite automaton whose states represent truth assignments of the chosen predicates.
    - The abstract model is checked for the desired property.
      - If the abstract model is error-free, the system is correct; otherwise an abstract counterexample is produced.
      - If it is checked whether the abstract counterexample corresponds to a real counterexample; if yes, the system is not correct.
      - If not, the chosen set of predicates contains too little information to verify or falsify the program; new predicates are added to the set.
        - Then the process is repeated.
  - Core problem: how to refine the abstraction.
    - Automated theorem provers are applied here.

Various model checkers for software verification use this approach.