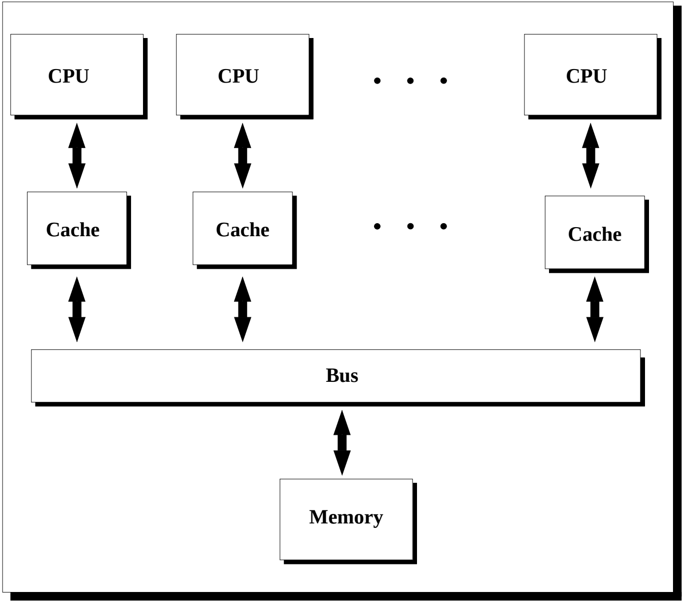
Parameterized Systems

- **Models**
  Families of finite-state machines indexed on N=number of processese
  Checking safety properties=*parameterized* reachability
- **Application**
  Consistency protocols designed for *multiprocessors* systems with *local caches*

CC-UMA Multiprocessor Systems



Cache Coherence Protocol

- **Goal**
  To ensure the consistency of the data stored in caches and main memory
- **Specification**
  Behavior of a single cache on **read**/**write** commands from **Bus**/**CPU**
- **Assumptions**
  Caches have a finite number of possible states
  They all behave *identically*
  We consider single cache lines

Formal Model for Protocols

A protocol is a tuple $P = \langle Q, \Sigma, \overline{\Sigma}, \tau \rangle$ where

- Q = cache states
- Σ = CPU commands
- $\overline{\Sigma}$ = Bus commands
- τ = transition relation, totally defined over $\overline{\Sigma}$

Global Machine with n processors

- **Global state**

$$\langle s_1, \ldots, s_n \rangle \in Q^n$$

- **Transition relation**

$$\tau_{\mathcal{M}}(\langle s_1, \ldots, s_n \rangle, \sigma) = \langle s'_1, \ldots, s'_n \rangle$$
$$\textit{if and only if}$$
$$\tau(s_i, \sigma) = s'_i \;\; \textit{and for all} \;\; j \neq i \;\; \tau(s_j, \overline{\sigma}) = s'_j$$

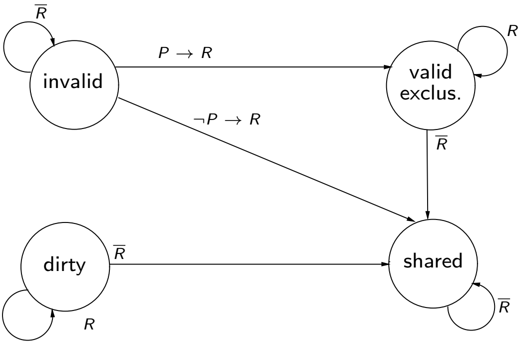
Global Conditions

To specify coherence policies we need actions 'guarded' by predicates

$$P \quad ::= \quad P \wedge P \mid P \vee P \mid \#q = c \mid \#q \geq c \mid \mathit{true}$$

$\#q \; = \;$ *number* of caches in state $q \in Q$ in the current global state

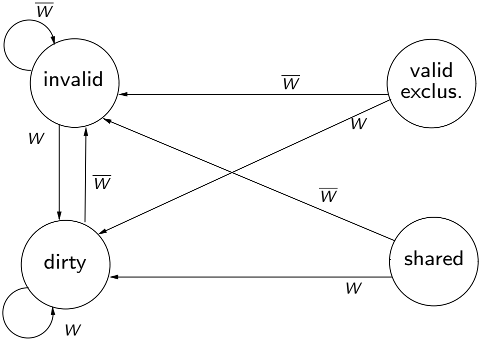
University of Illinois Protocol: Read Cycle



$$R = \text{read cache}$$

$$P \equiv \#dirty = 0 \wedge \#shared = 0 \wedge \#valid = 0$$

University of Illinois Protocol: Write Cycle



W = write in cache

Sample Run for n=3

$$\begin{array}{cl}
\overset{mem=fresh}{\langle inv, inv, inv\rangle} & \\
\downarrow & cpu_1 : R \\
\overset{mem=fresh}{\langle valid_excl, inv, inv\rangle} & \\
\downarrow & cpu_2 : R,\ \ bus : \overline{R} \\
\overset{mem=fresh}{\langle shared, shared, inv\rangle} & \\
\downarrow & cpu_2 : W,\ \ bus : \overline{W} \\
\overset{mem=obsolete}{\langle inv, dirty, inv\rangle} & \\
\downarrow & cpu_2 : Rep \\
\overset{mem=fresh}{\langle inv, inv, inv\rangle} &
\end{array}$$

Safety Properties

- **Data consistency**
  In every reachable global state there is at most one *dirty* cache;
  furthermore, *dirty* and *shared* caches cannot coexist
- **Parameterized reachability problem**
  A safety property is violated whenever there exists N such that an *unsafe* state is reachable in the global machine with N processors

Counting Abstraction

$$\mathbf{G} = \langle s_1, \ldots, s_n \rangle \quad \longrightarrow \quad \mathbf{G}^{\#} = \langle Occ_{q_1}(G), \ldots, Occ_{q_K}(G) \rangle$$

$Occ_q(G)$=*number* of occurrences of $q \in Q$ in G

$$\langle shared, shared, invalid \rangle \quad \longrightarrow \quad \langle 1, 2, 0, 0 \rangle$$

Abstract Protocol = Extended Finite-state Machine (EFSM)

Transition $\longrightarrow$ **Guarded command** over *integer counters*

$$\tau(invalid, R) = valid \;\; if \; \#valid = 0$$
$$becomes$$
$$x_{invalid} \geq 1, x_{valid} = 0, x'_{invalid} = x_{invalid} - 1, x'_{valid} = x_{valid} + 1$$

Sample Abstract Run for n=3

$$
\begin{array}{ccc}
\overset{mem=fresh}{\langle inv, inv, inv \rangle} & & \langle 3,0,0,0 \rangle \\
\downarrow & cpu_1 : R & \downarrow \\
\overset{mem=fresh}{\langle valid\_excl, inv, inv \rangle} & & \langle 2,1,0,0 \rangle \\
\downarrow & cpu_2 : R,\ \ bus : \overline{R} & \downarrow \\
\overset{mem=fresh}{\langle shared, shared, inv \rangle} & & \langle 1,0,2,0 \rangle \\
\downarrow & cpu_2 : W,\ \ bus : \overline{W} & \downarrow \\
\overset{mem=obsolete}{\langle inv, dirty, inv \rangle} & & \langle 2,0,0,1 \rangle \\
\downarrow & cpu_2 : Rep & \downarrow \\
\overset{mem=fresh}{\langle inv, inv, inv \rangle} & & \langle 3,0,0,0 \rangle
\end{array}
$$

Verification = EFSM Reachability

- **Initial states**
  $\Phi_I \;=\; x_{invalid} \geq 0, x_{dirty} = 0, x_{shared} = 0, x_{valid} = 0$
- **Unsafe states** $\Phi_U \;=\; x_{dirty} \geq 2 \quad \vee \quad x_{dirty} \geq 1, x_{shared} \geq 1$
- **Reachability = Full Test** The protocol is safe *iff* Φ_U is not EFSM-*reachable* from Φ_I

Symbolic Model Checking

- **Symbolic Representation = Integer Constraints**

$$
\begin{aligned}
[\![x_{invalid} \geq 2]\!] &= \{\langle 2, 0, \ldots \rangle, \langle 3, 1, \ldots \rangle, \ldots\} \\
&= \{\langle \mathit{invalid}, \mathit{invalid} \rangle, \\
&\quad \langle \mathit{invalid}, \mathit{shared}, \mathit{invalid} \rangle, \ldots\}
\end{aligned}
$$

- **Entailment Test**

$$\varphi \sqsubseteq \psi \quad \textit{if and only if} \quad [\![\psi]\!] \subseteq [\![\varphi]\!]$$

- **Symbolic Predecessor Operator**

$$\textbf{sym\_pre}(\varphi(\mathbf{x}')) = \bigvee_{i \in I} \; \exists\, \mathbf{x}'.\, \psi_\tau(\mathbf{x}, \mathbf{x}') \;\wedge\; \varphi(\mathbf{x}')$$

Decidable Issues

For generic guards

- Parameterized verification is undecidable: counter machines (i.e. with zero test) are a subclass of EFSM
- Backward reachability may not terminate, each step is effective: verification procedure (it may find bugs)

Decidable Subclass: L-constraints

Let $x_1, \ldots, x_n$ be variables over natural numbers Let us restrict our attention to L-constraints, i.e., conjunctions of atomic formulas of the following form

$$x_{i_1} + \ldots + x_{i_n} \geq c$$

where $x_l \neq x_m$ or $l \neq m$

A Decidable Subclass

- Guards are restricted to L-constraints (i.e. no test for zero/constants)
- Set of states are symbolically expressed via sets of L-constraints

Properties

- L-constraints represent upward closed set of tuples of natural numbers ordered via pointwise ordering
- L-constraints are closed under application of **sym_pre**
- L-constraints are always satisfiable
- checking containment of sets of L-constraints is co-NP complete
- entailment (i.e., given two L-constraints ϕ and ψ, does ϕ entail ψ?) is co-Np-complete

S-constraints

Conjunctions of atomic formulas of the form $x_i \geq c$

- they are not closed under application of **sym_pre**
- containment of sets of S-constraints is polynomial
- entailment is polynomial
- Ł-constraints can be reduced to sets of L-constraints $x_{i_1} + \ldots + x_{i_m} \geq c$ can be decomposed as follows:

$$\bigvee_{c_1+\ldots+c_m=c} x_{i_1} \geq c_1 \wedge x_{i_2} \geq c_2 \wedge \ldots \wedge x_{i_m} \geq c_m$$

Possible algorithms for model checking

- Keep constraints in S-normal form
  Entailment and containment: polynomial in size of sets and constraints
  Size of intermediate results: each step exponential explosion
- Keep constraints in L-form
  Entailment: polynomial in size of constraints
  Size of Intermediate results: each step polynomial (in the constants)
  Replace 'full containment test' (in co-NP) with 'local containment (in P)'

Termination

- For EFSMs in which guards are L-constraints, symbolic backward reachability with pointwise entailment terminates
- Indeed, the entailment relation over L- and S-constraints is a wqo
- This follows from Dickson's lemma and by composition properties of wqo's