Ingeniería del Software II 6 - Model Checking

El problema de model checking

Dado un modelo M de un sistema (en algún lenguaje) y una propiedad ϕ (en alguna lógica) deseamos verificar automáticamente si esta es satisfecha por M, es decir, si $M \models \phi$.

En particular para el caso de LTL:

- sabemos que el lenguaje $\mathcal{L}(\phi)$ de una fórmula es el conjunto de todas las trazas donde ésta se hace verdadera, y que
- el comportamiento de un sistema M (denotado $\mathcal{L}(M)$) está dado por el conjunto de todas las trazas que éste puede ejecutar.

Luego, $M \models \phi$ si y sólo si toda traza de M satisface ϕ , es decir:

$$M \models \phi$$
 si y sólo si $\mathcal{L}(M) \subseteq \mathcal{L}(\phi)$

El problema de model checking se reduce entonces a validar esta inclusión de manera automática

¿Pero cómo?

Cómo obtener el modelo de un sistema

Para poder definir $\mathcal{L}(M)$ necesitamos primero una manera razonable de definir M. Pero eso ya sabemos como hacerlo:

```
int y1 = 0;
  int y2 = 0;
  short in_critical = 0;
  active proctype process_1() {
                                        active proctype process_2() {
    do
                                          do
                                          :: true ->
    :: true ->
         y1 = y2+1;
                                            y2 = y1+1;
0:
                                      0:
         ((y2==0) | | (y1<=y2));
                                              ((y1==0) | | (y2 < y1));
1:
                                                                                (0,0,0,0,0)
                                      1:
         in_critical++;
                                               in_critical++;
         in_critical--;
                                               in_critical--;
2:
                                      2:
                                                                        (1,0,1,0,0)
                                                                                       (0, 1, 0, 1, 0)
         v1 = 0;
                                               y2 = 0;
3:
                                      3:
                                          od
    od
                                                            (1,1,1,2,0) (2,0,1,0,1)
                                                                                       (0,2,0,1,1) (1,1,2,1,0)
                                                            (2,1,1,2,1) (3,0,1,0,0)
                                                                                       (0,3,0,1,0) (1,2,2,1,1)
 Estructura del estado:
                                                            (3, 1, 1, 2, 0)
                                                                                                   (1,3,2,1,0)
    (pc_1, pc_2, y1, y2, in\_critical)
                                                            (0, 1, 0, 2, 0)
                                                                                                   (1,0,2,0,0)
```

Cómo obtener el modelo de un sistema

Es decir, el modelo del sistema define un sistema de transiciones.

Un sistema de transiciones es una estructura

$$M = (S, s_0, \rightarrow, v)$$

donde:

- S es un conjunto de **estados** donde $s_0 \in S$ es el **estado inicial**,
- $\rightarrow \subseteq S \times S$ es la **relación de transición** tal que $\forall s \in S : \exists s' \in S : s \rightarrow s'$
- $v: S \to 2^{PA}$ es una función de **valuación**.

v(s) es el conjunto de todas las proposiciones atómicas que son verdaderas en el estado s.

(En el contexto de lógicas modales, esta estructura se denomina estructura de Kripke)

Cómo obtener el modelo de un sistema

Una ejecución de M es una función $\rho: \mathbb{N} \to S$ tal que:

1.
$$\rho(0) = s_0$$
, y

2.
$$\rho(i) \rightarrow \rho(i+1)$$
 para todo $i \geq 0$

El comportamiento de M se define como:

$$\mathcal{L}(M) = \{ \sigma \in \left(2^{PA}\right)^{\omega} \mid \exists \rho \text{ ejecución de } M: \ \forall i \geq 0: \sigma(i) = v(\rho(i)) \}$$

Autómatas de Büchi

Al igual que los lenguajes regulares, los lenguajes ω -regulares no son fáciles de manipular por sí mismos.

De la misma manera que los lenguajes regulares se manipulan a través de autómatas finitos que acepten los lenguajes a manipular, los lenguajes ω -regulares pueden manipularse a través de los denominados autómatas de Büchi.

Un autómata de Büchi es una estructura

$$\mathcal{A} = (\Sigma, S, \delta, s_0, A)$$

en la cual:

- Σ es un conjunto finito, llamado alfabeto,
- S es un conjunto finito de estados donde $s_0 \in S$ es el estado inicial,
- $\delta: \Sigma \times S \to 2^S$ es la función de transición,
- A es el conjunto de estados de aceptación.

Aceptación de trazas en autómatas de Büchi

Dado un autómata de Buchi \mathcal{A} , decimos que una traza $\sigma = \sigma_0$ $\sigma_1 \sigma_2 \dots de$ elementos de Σ es aceptada por \mathcal{A} sii existe una función $\rho: IN \to S$ (denominada ejecución) tal que:

- $\rho(0) = s_0$,
- $\rho(i+1) \in \delta(\sigma(i), \rho(i))$ para todo $i \geq 0$, y
- existen estados de aceptación en A que se repiten infinitas veces en ρ , i.e., el conjunto $\{i \in I\!\!N \mid \rho(i) \in A\}$ es infinito.

Definimos como el lenguaje de \mathcal{A} , notación $\mathcal{L}(\mathcal{A})$, al conjunto de todas las trazas (i.e., ω -palabras) aceptadas por \mathcal{A} .

Los autómatas de Büchi aceptan exactamente todos los lenguajes ω -regulares. Por consiguiente son más expresivos que LTL sobre el alfabeto 2^{PA} .

Autómatas de Büchi como modelos de sistemas

Ya dijimos que un programa P cuyo espacio de estado sea finito puede representarse con un sistema de transiciones finito $M_P = (S, s_0, \rightarrow, v)$.

A su vez, M_P puede verse como el autómata de Büchi $\mathcal{A}_P = (\Sigma, S, \delta, s_0, S)$, donde:

- \bullet $\Sigma = 2^{PA}$,
- $s_j \in \delta(B, s_i)$ sii $s_i \to s_j \land B = v(s_i)$

Notar que **todos** los estados son estados de aceptación. Esto es así porque nos interesan todas las ejecuciones posibles del sistema.

Teorema:
$$\mathcal{L}(M_P) = \mathcal{L}(\mathcal{A}_P)$$

Autómatas de Büchi como modelos de sistemas

Ya dijimos que un programa P cuyo espacio de estado sea finito puede representarse con un sistema de transiciones finito $M_P = (S, s_0, \rightarrow, v)$.

A su vez, M_P puede verse como el autómata de Büchi $\mathcal{A}_P = (\Sigma, S, \delta, s_0, S)$, donde:

- $\bullet \ \Sigma = 2^{PA},$
- $s_j \in \delta(B, s_i)$ sii $s_i \to s_j \land B = v(s_i)$

Notar que todos los estados son estados de aceptación.

Esto es así porque nos interesan todas la posibles del sistema.

Demostrar formalmente este teorema.

Teorema:
$$\mathcal{L}(M_P) = \mathcal{L}(\mathcal{A}_P)$$

Fórmulas LTL y Autómatas de Büchi

Teorema: Para toda fórmula LTL ϕ , se puede construir un autómata de Büchi \mathcal{A}_{ϕ} tal que:

$$\mathcal{L}(\mathcal{A}_{\phi}) = \mathcal{L}(\phi)$$

La demostración de este teorema es compleja. Para formarse una idea de lo establecido por el teorema daremos algunos ejemplos:

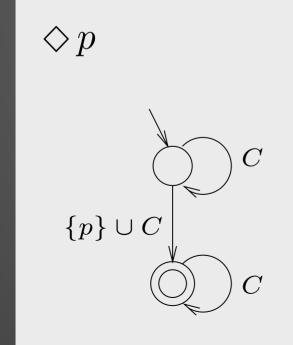


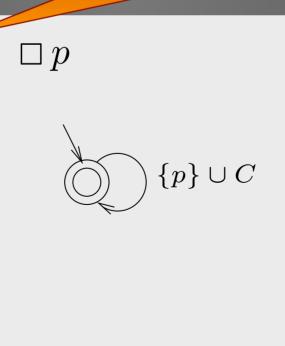
Fórmulas LTL y Autómatas de Büchi

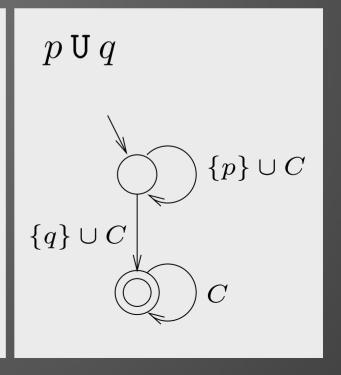
Teorema: Para toda fórmula LTL ϕ , se puede construir un autómata de Büchi \mathcal{A}_{ϕ} tal que:

$$\mathcal{L}(\mathcal{A}_{\phi}) = \mathcal{L}(\phi)$$

La demostración de este teorema idea de lo establecido por el te C es cualquier subconjunto de \mathcal{PA} . Es decir, cada flecha de los dibujos representa muchas transiciones a la vez.







Manipulación de lenguajes ω-regulares usando autómatas de Büchi

Teorema: Dados dos autómatas de Büchi \mathcal{A}_1 y \mathcal{A}_2 , se puede construir un automata $\mathcal{A}_1 \cap \mathcal{A}_2$ tal que:

$$\mathcal{L}(\mathcal{A}_1 \cap \mathcal{A}_2) = \mathcal{L}(\mathcal{A}_1) \cap \mathcal{L}(\mathcal{A}_2)$$

Teorema: Dado un autómada de Büchi \mathcal{A} se puede construir un automata \mathcal{A}^c tal que:

$$\mathcal{L}(\mathcal{A}^c) = \overline{\mathcal{L}(\mathcal{A})}$$

Teorema: Existe un algoritmo que permite decidir si el lenguaje ω -regular aceptado por un autómata de Büchi es vacío o no.

Manipulación de lenguajes ω-regulares usando autómatas de Büchi

Teorema: Dados dos autómatas de Büchi \mathcal{A}_1 y \mathcal{A}_2 , se puede construir un automata $\mathcal{A}_1 \cap \mathcal{A}_2$ tal que:

$$\mathcal{L}(\mathcal{A}_1 \cap \mathcal{A}_2) = \mathcal{L}(\mathcal{A}_1) \cap \mathcal{L}(\mathcal{A}_2)$$

Teorema: Dado un autómada dun automata \mathcal{A}^c tal que:

$$\mathcal{L}(\mathcal{A}^c) = \overline{\mathcal{L}(\mathcal{A})}$$

El algoritmo es un doble

DFS con el fin de buscar componentes
fuertemente conexas que atrapen un
estado de aceptación.

Teorema: Existe un algoritmo que permite decidir si el lenguaje ω -regular aceptado por un autómata de Büchi es vacío o no.

Model checking con fundamentos en la teoría de autómatas

Hemos visto que la verificación de que un programa P de estados finitos satisfaga una propiedad temporal ϕ (i.e., $M_P \models \phi$) se reduce a comprobar que

$$\mathcal{L}(M_P) \subseteq \mathcal{L}(\phi)$$

Con los resultados anteriores, podemos reducir este problema a verificar si

$$\mathcal{L}(\mathcal{A}_P) \subseteq \mathcal{L}(\mathcal{A}_\phi)$$

que a su vez es equivalente a verificar si:

$$\mathcal{L}(\mathcal{A}_P) \cap \overline{\mathcal{L}(\mathcal{A}_\phi)} = \emptyset$$

Problema: Complementar un autómata de Büchi es computacionalmente muy caro (se produce una explosión exponencial).

¿Cómo podemos evitar este problema?

Model checking con fundamentos en la teoría de autómatas

Hemos visto que la verificación de que un programa P de estados finitos satisfaga una propiedad temporal ϕ (i.e., $M_P \models \phi$) se reduce a comprobar que

$$\mathcal{L}(M_P) \subseteq \mathcal{L}(\phi)$$

Con los resultados anteriores, podemos reducir este problema a verificar si

$$\mathcal{L}(\mathcal{A}_P) \subseteq \mathcal{L}(\mathcal{A}_\phi)$$

que a su vez es equivalente a verificar si:

$$\mathcal{L}(\mathcal{A}_P) \cap \overline{\mathcal{L}(\mathcal{A}_\phi)} = \emptyset$$

Problema: Complementar un autómata de Büchi es computacionalmente muy caro (se produce una explosión exponencial). Por suerte:

$$\overline{\mathcal{L}(\mathcal{A}_{\phi})} = \mathcal{L}(\mathcal{A}_{\neg \phi})$$

Por lo tanto, podemos verificar equivalentemente que:

$$\mathcal{L}(\mathcal{A}_P) \cap \mathcal{L}(\mathcal{A}_{\neg \phi}) = \emptyset$$

El algoritmo de model checking "in a nutshell"

Aplicando los resultados anteriores el problema de verificar si un programa P satisface una propiedad ϕ puede esquematizarse como sigue:

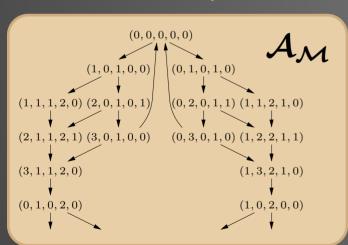
- 1. Construir el autómata de Büchi \mathcal{A}_P
- 2. Construir el autómata de Büchi $\mathcal{A}_{\neg\phi}$
- 3. Construir el autómata de Büchi $\mathcal{A}_P \cap \mathcal{A}_{\neg \phi}$
- 4. Comprobar si $\mathcal{L}(\mathcal{A}_P \cap \mathcal{A}_{\neg \phi})$ es vacío

Un aspecto muy importante del model checking (sino el más importante) es la obtención de un contraejemplo en caso de que la propiedad no sea verdadera: ¿Cómo se obtiene tal contraejemplo?

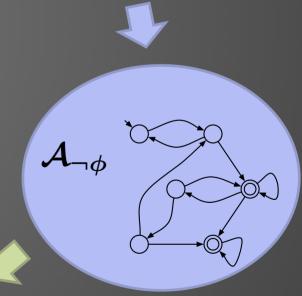
```
\mathcal{M} \models \phi?
```

```
int v1 = 0;
int y2 = 0;
short in_critical = 0;
active proctype process_1() {
                                    active proctype process_2() {
  :: true ->
                                      :: true ->
      y1 = y2+1;
                                           y2 = y1+1;
       ((y2==0) | | (y1<=y2));
                                           ((y1==0) || (y2<y1));
       in_critical++;
                                           in_critical++;
      in_critical--;
                                           in_critical--;
                                           y2 = 0;
       y1 = 0;
 od
                                      od
```

```
\phi: \Box \diamondsuit crit_1 \land \Box \diamondsuit crit_2
\neg \phi: \neg (\Box \diamondsuit crit_1 \land \Box \diamondsuit crit_2)
```



El problema de model checking (para LTL)



$$i \mathcal{A}_{\mathcal{M}} \cap \mathcal{A}_{\neg \phi} = \emptyset$$
?

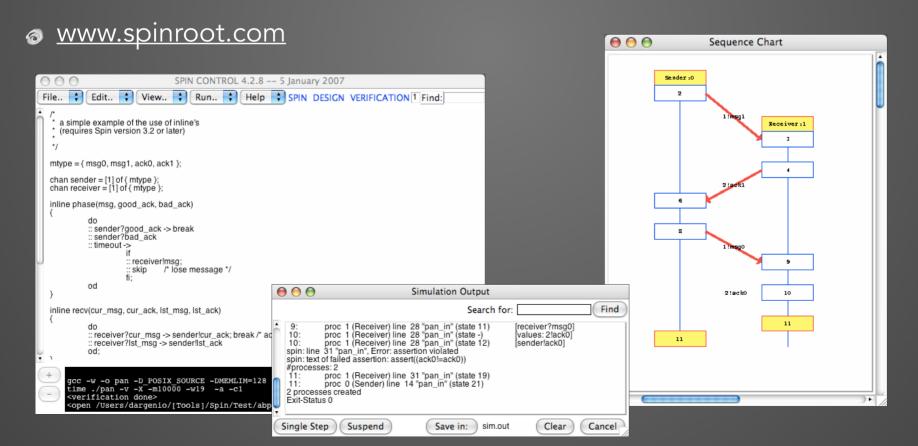
Model Checking: características

- Además de determinar si una propiedad se cumple, dan contraejemplos en caso de que no se cumpla.
- El algoritmo básico se basa en "fuerza bruta": recorre todo el grafo subyacente.
- Esto se agrava con el problema de la explosión de estados.
 - Se agranda exponencialmente con cada variable y cada proceso.
- El grafo subyacente usualmente necesita ser finito.

Herramientas de Model Checking

El model checker SPIN

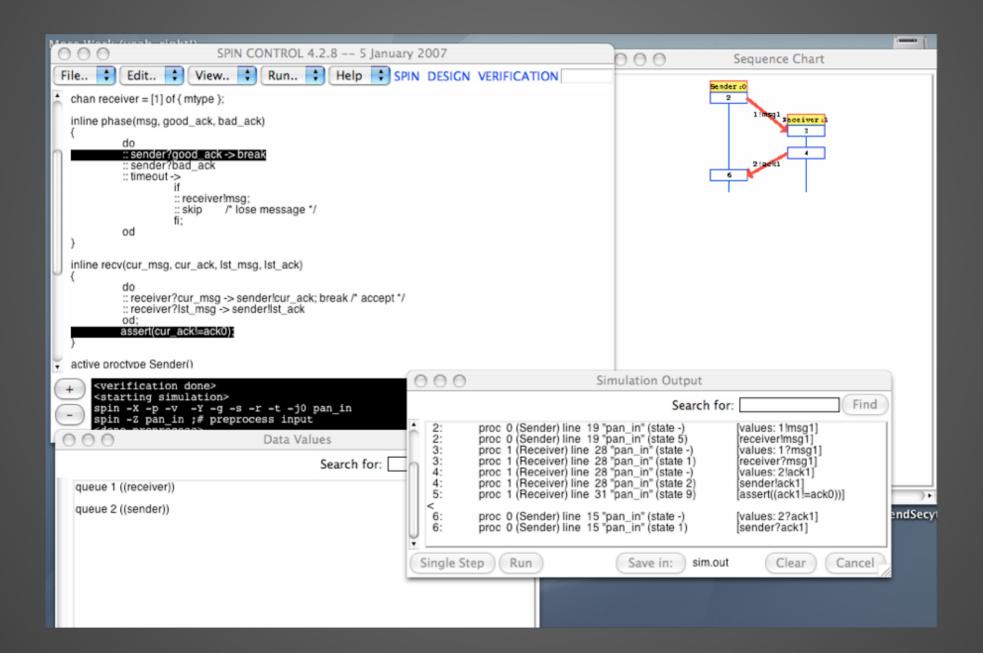
- Desarrollado en AT&T / Bell Labs.
- Principalmente desarrollado por Gerard Holzmann
- Bibliografía:
 - G. Holzmann. The Spin Model Checker. Addisson Wesley. 2004.



- La descripción de los modelos se realiza en PROMELA
- Promela se asemeja a C y agrega primitivas para manejar concurrencia, canales, atomicidad, no determinismo, ...

```
/*
                                                   inline recv(cur msg, cur ack, lst msg, lst ack)
    The alternating bit protocol.
                                                      do
    A simple example of the use of inline's
                                                       :: receiver?cur msg -> sender!cur ack;
 */
                                                                               break /* accept */
                                                       :: receiver?lst msg -> sender!lst ack
mtype = { msq0, msq1, ack0, ack1 };
                                                      od;
                                                   }
chan sender = [1] of { mtype };
chan receiver = [1] of { mtype };
                                                   active proctype Sender()
                                                      do
inline phase (msg, good ack, bad ack)
                                                       :: phase (msq1, ack1, ack0);
                                                         phase (msq0, ack0, ack1)
   do
                                                      od
    :: sender?good ack -> break
                                                   }
    :: sender?bad ack
    :: timeout ->
                                                   active proctype Receiver()
       if
                                                      do
       :: receiver!msg;
                                                       :: recv(msq1, ack1, msq0, ack0);
       :: skip/* lose message */
                                                         recv(msg0, ack0, msg1, ack1)
       fi;
                                                      od
    od
}
```

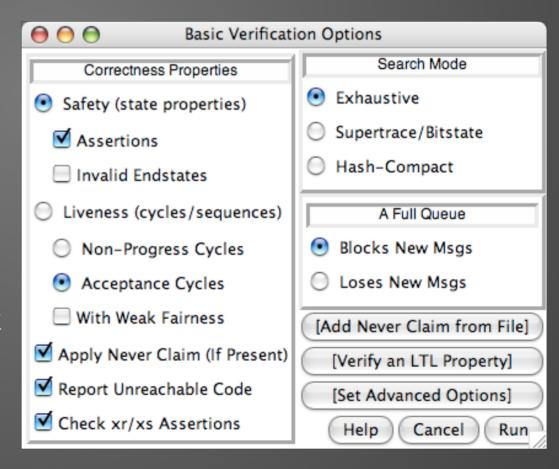
Permite realizar simulaciones guiadas, aleatorias, sobre una traza específica (ej: contraejemplo de una propiedad).



El model checker SPIN (cont.)

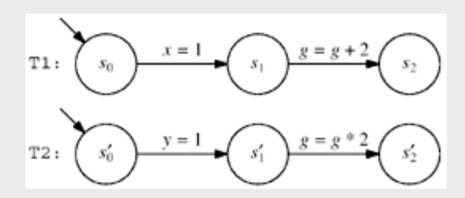
Permite distintos tipos de verificaciones:

- Propiedades en LTL
- Aserciones dentro del modelo
- Deadlocks
- Progreso
- Permite verificar bajo weak fairness

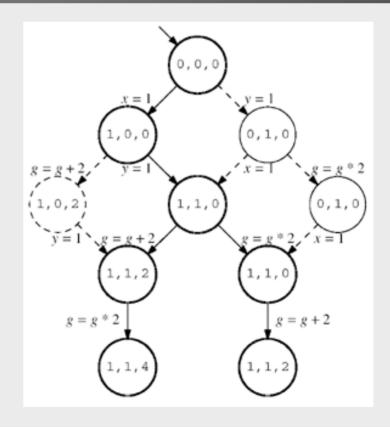


El model checker SPIN (cont.) Técnicas de optimización

- **Bitstate hashing**: los visitados en el DFS se marcan usando una tabla hash con imagen en {0,1}.
- Reducción por orden parcial: aprovecha la simetría introducida por el "interleaving":



Más información y más técnicas y algoritmos en el Cap. 9 de "The Model Checker Spin"



El model checker SPIN (cont.) Usos

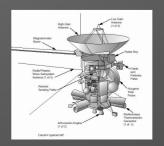
- Spin se ha utilizado en múltiples ocasiones, y en particular, directamente en la industria (¡se implementó en la industrial!).
- Además es utilizado en la academia para aplicaciones reales (subcontratos/proyectos por parte de empresas).

Ejemplos:

- Verificación de protocolos embebidos en automotores (Bosch)
- Verificación del dique de emergencia climática en Rotterdam.
- Software para el procesamiento de llamadas (Lucent Tech.)
- Diversos algoritmos en proyectos de la NASA como Deep Space 1, Cassini, Mars Exploration Rovers, Deep Impact, etc.







El model checker SMV

- SMV fue originalmente desarrollado por Ken McMillan / Edmund Clarke en Carnegie-Mellon University.
- El SMV original derivó en múltiples versiones:
 - SMV CMU (www.cs.cmu.edu/~modelcheck/smv.html)
 - SMV Cadence (inaccessible)
 - NuSMV (nusmv.fbk.eu) / nuXmv (nuxmv.fbk.eu)
 - => Elegir éste! (LGPL, más nuevo, único mantenido)
- Originalmente destinado a la verificación de hardware.
- Uso en línea de comandos :-(
- Manipula espacio de estados enormes.

- El lenguaje de SMV es bastante básico.
- Describe redes de autómatas (con composición sincrónica o asincrónica según se especifique).
- Descripción de cada autómata bastante declarativa, usando variables y un predicado "next" que permite hablar del valor de las variables en el siguiente estado.
- Simplemente de esa manera se definen las transiciones.

```
MODULE main
VAR
  semaphore : boolean;
  proc1 : process user(semaphore);
  proc2 : process user(semaphore);
ASSTON
  init(semaphore) := 0;
SPEC
  AG (proc1.state = entering
        -> AF proc1.state = critical)
MODULE user (semaphore)
VAR
  state : {idle,entering,critical,exiting};
ASSTON
  init(state) := idle;
  next(state) :=
    case
      state = idle : {idle,entering};
      state = entering & !semaphore : critical;
      state = critical : {critical,exiting};
      state = exiting : idle;
      1 : state;
    esac:
  next(semaphore) :=
    case
      state = entering : 1;
      state = exiting : 0;
      1 : semaphore;
    esac:
FAIRNESS
  running
```

```
MODULE main
VAR
  semaphore : boolean;
  proc1 : process user(semaphore);
  proc2 : process user(semaphore);
ASSTGN
  init(semaphore) := 0;
SPEC
  AG (proc1.state = entering
        -> AF proc1.state = critical)
MODULE user (semaphore)
VAR
  state : {idle,entering,critical,exiting};
ASSIGN
  init(state) := idle;
  next(state) :=
    case
      state = idle : {idle,entering};
      state = entering & !semaphore : critical;
      state = critical : {critical,exiting};
      state = exiting : idle;
      1 : state;
    esac;
  next(semaphore) :=
    case
      state = entering : 1;
      state = exiting : 0;
      1 : semaphore;
    esac;
FAIRNESS
  running
```

Simulación es posible pero a través de la línea de comandos

```
MODULE main
VAR
  semaphore : boolean;
  proc1 : process user(semaphore);
  proc2 : process user(semaphore);
ASSIGN
  init(semaphore) := 0;
SPEC
  AG (proc1.state = entering
        -> AF proc1.state = critical)
MODULE user (semaphore)
VAR
  state : {idle,entering,critical,exiting};
ASSIGN
  init(state) := idle;
  next(state) :=
    case
      state = idle : {idle,entering};
      state = entering & !semaphore : critical;
      state = critical : {critical,exiting};
      state = exiting : idle;
      1 : state;
    esac;
  next(semaphore) :=
    case
      state = entering : 1;
      state = exiting : 0;
      1 : semaphore;
    esac;
FAIRNESS
  running
```

Simulación es yés de la Las propiedades se expresan usando las lógicas CTL, LTL, o PSL

```
MODULE main
VAR
  semaphore : boolean;
  proc1 : process user(semaphore);
  proc2 : process user(semaphore);
ASSTGN
  init(semaphore) := 0;
SPEC
  AG (proc1.state = entering
        -> AF proc1.state = critical)
MODULE user (semaphore)
VAR
  state : {idle,entering,critical,exiting};
ASSIGN
  init(state) := idle;
  next(state) :=
    case
      state = idle : {idle,entering};
      state = entering & !semaphore : critical;
      state = critical : {critical,exiting};
      state = exiting : idle;
      1 : state:
    esac;
  next(semaphore) :=
    case
      state = entering : 1;
      state = exiting : 0;
      1 : semaphore;
    esac;
FATRNESS
  running
```

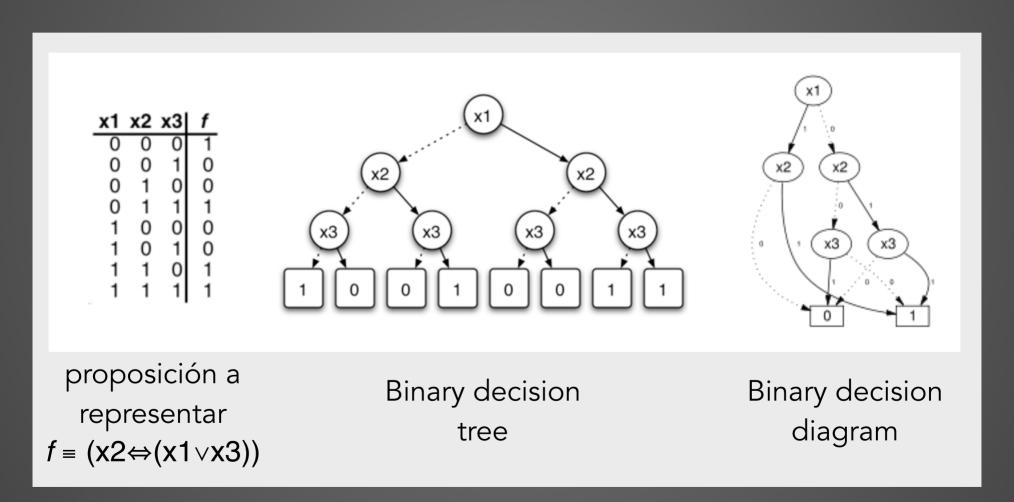
Simulación es

Las propiedades se
expresan usando las lógicas
CTL, LTL, o PSL

Es posible especificar que se desea hacer la verificación bajo la suposición de fairness

El model checker (Nu)SMV (cont.) Técnicas de optimización

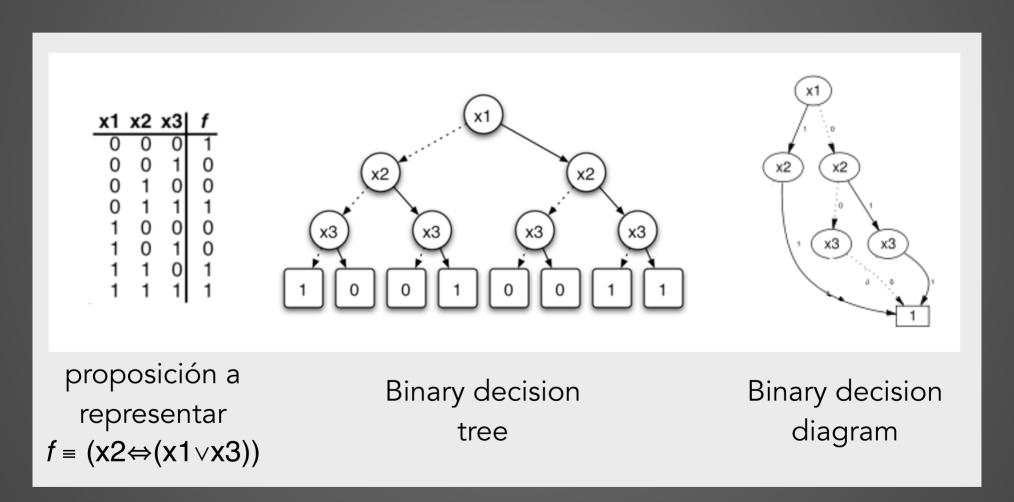
Representación del espacio de estado usando BDDs



Además: Bounded model checking utilizando SAT solvers.

El model checker (Nu)SMV (cont.) Técnicas de optimización

Representación del espacio de estado usando BDDs



Además: Bounded model checking utilizando SAT solvers.

El model checker (Nu)SMV (cont.) Usos

- Desde CMU se proveen servicios a:
 - National Science Foundation (NSF); Gigascale Systems Research Center (GSRC); Office of Naval Research (ONR); Army Research Office (ARO); Semiconductor Research Corporation (SRC); General Motors (GM).
- SMV se ha utilizado en múltiples ocaciones pero siempre desde la academia como servicio a la industria. Ej.:
 - Diversos protocolos para coherencia de cache (Gigamax, Futurebus+, etc.)
 - Diversos circuitos lógicos, componentes de procesadores y protocolos
 - Desafortunadamente no se reporta mucho en la literatura.