An Introduction to Probabilistic Model Checking

Pedro R. D'Argenio

Dependable Systems Group - FaMAF

Universidad Nacional de Córdoba
CONICET

http://dsg.famaf.unc.edu.ar

11º ERPEM, Rio Cuarto, Dec-2015

CONICET

UNC

Overview

- Motivation
- Reachability analysis on deterministic models
- Reachability analysis on non-deterministic models
- LTL
- The process of probabilistic model checking
- Quick and partial overview of the state of the art

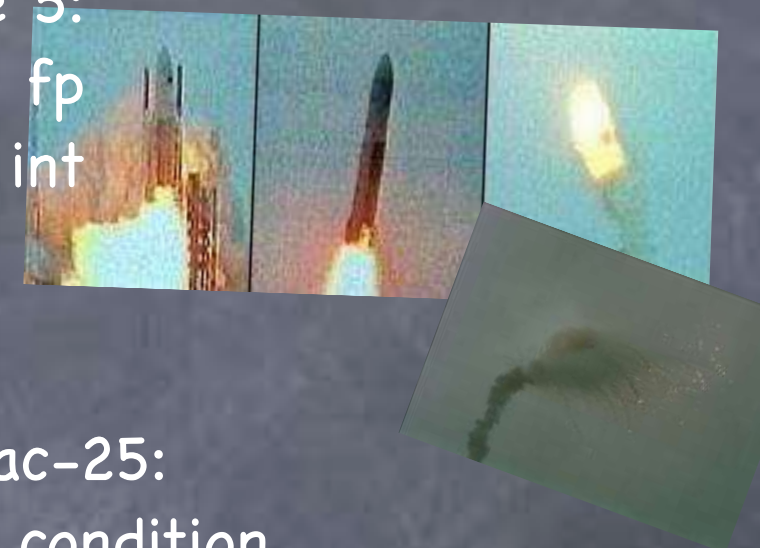
CONICET

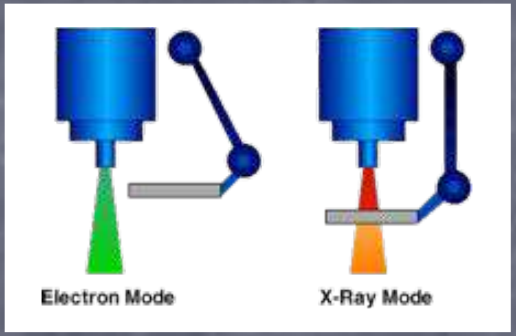
UNC

Why verification?

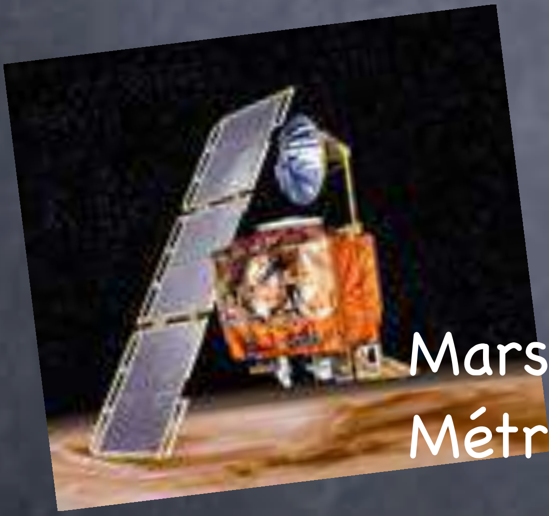
Pentium:
FDIV

Ariane 5:
64 bits fp
vs 16 bits int



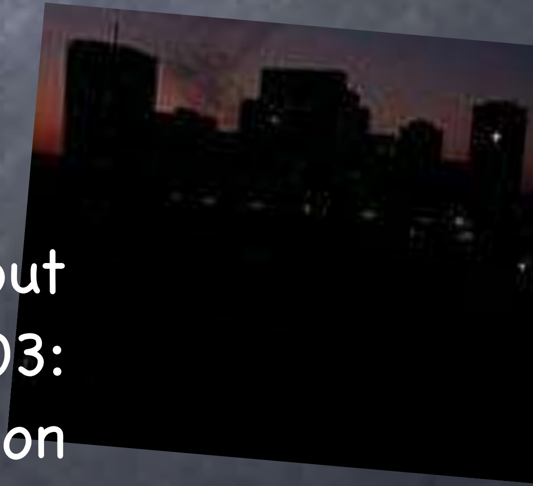


Therac-25:
Race condition



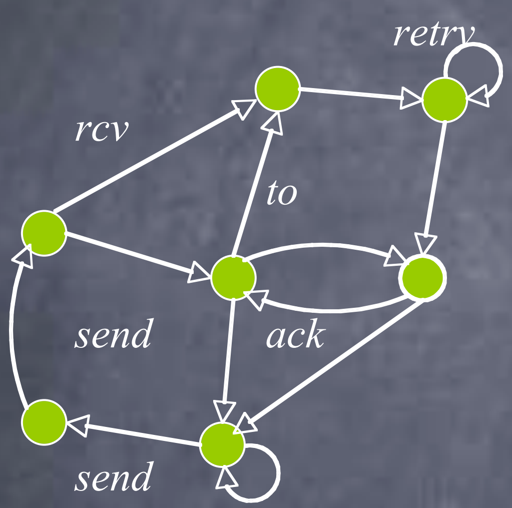
Mars Climate Orbiter:
Métrico vs Imperial

Northeast blackout
in 2003:
Race condition



Heartbleed:
Integridad/Confidencialidad

Model Checking



Properties are either true or false

$\models$ G (send(msg) => F rcv(msg))

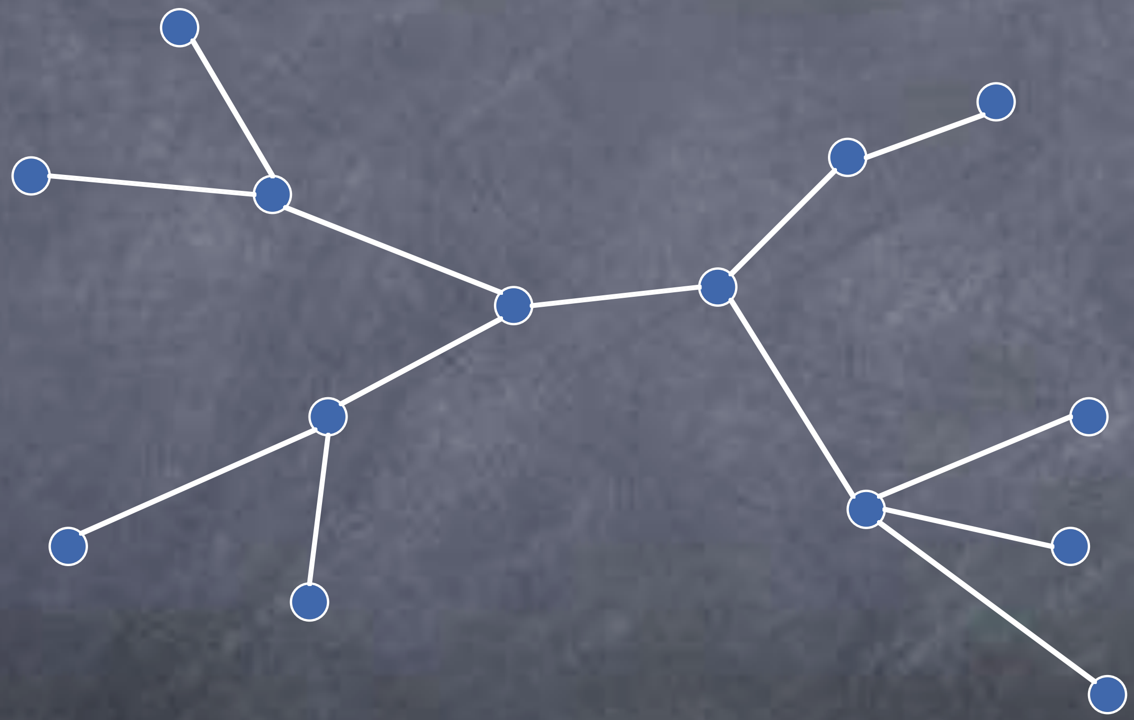
Non-deterministic behavior

Limitations of this approach

- Many algorithms proposed (better) solutions using randomization.
- E.g.
  - Leader election protocol in IEEE 1394 “Firewire”
  - Binary exponential backoff on IEEE 802.3 “Ethernet”

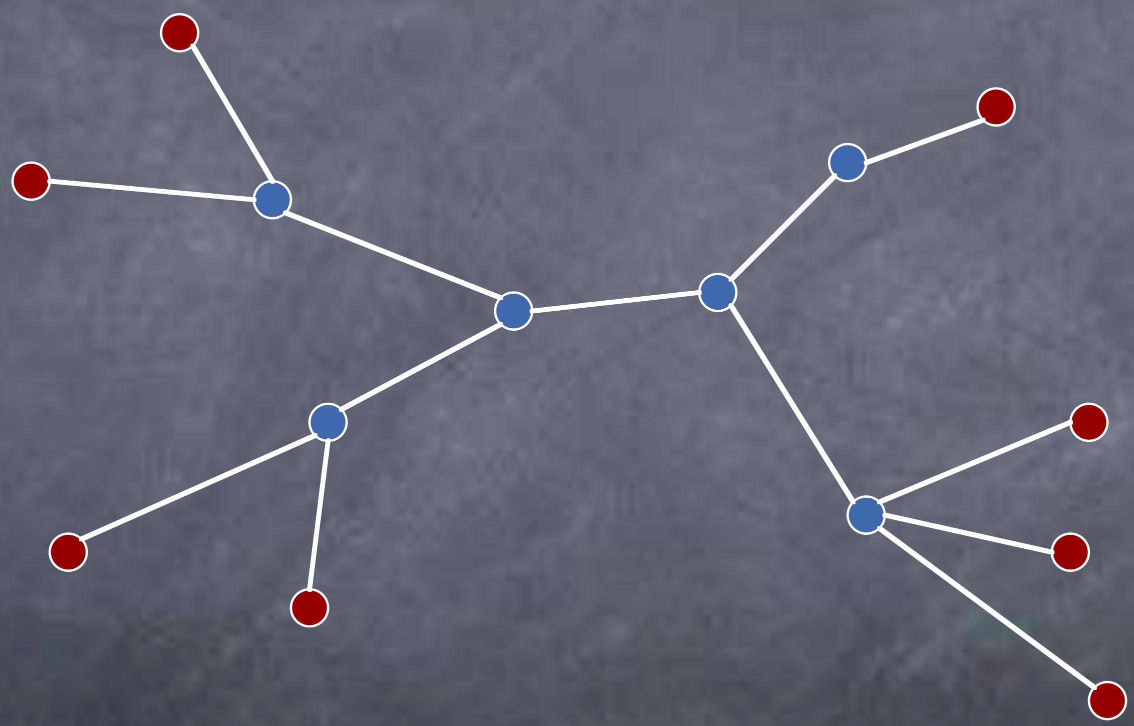
Limitations of this approach

E.g.: IEEE 1394 Leader election protocol



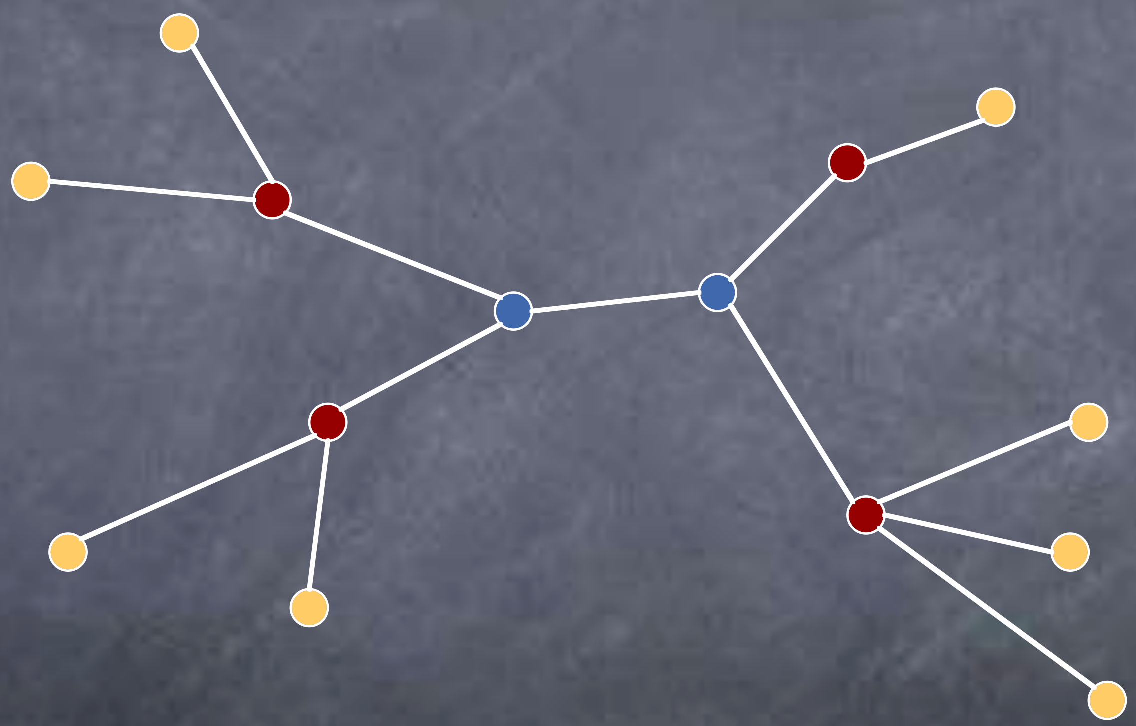
Limitations of this approach

E.g.: IEEE 1394 Leader election protocol



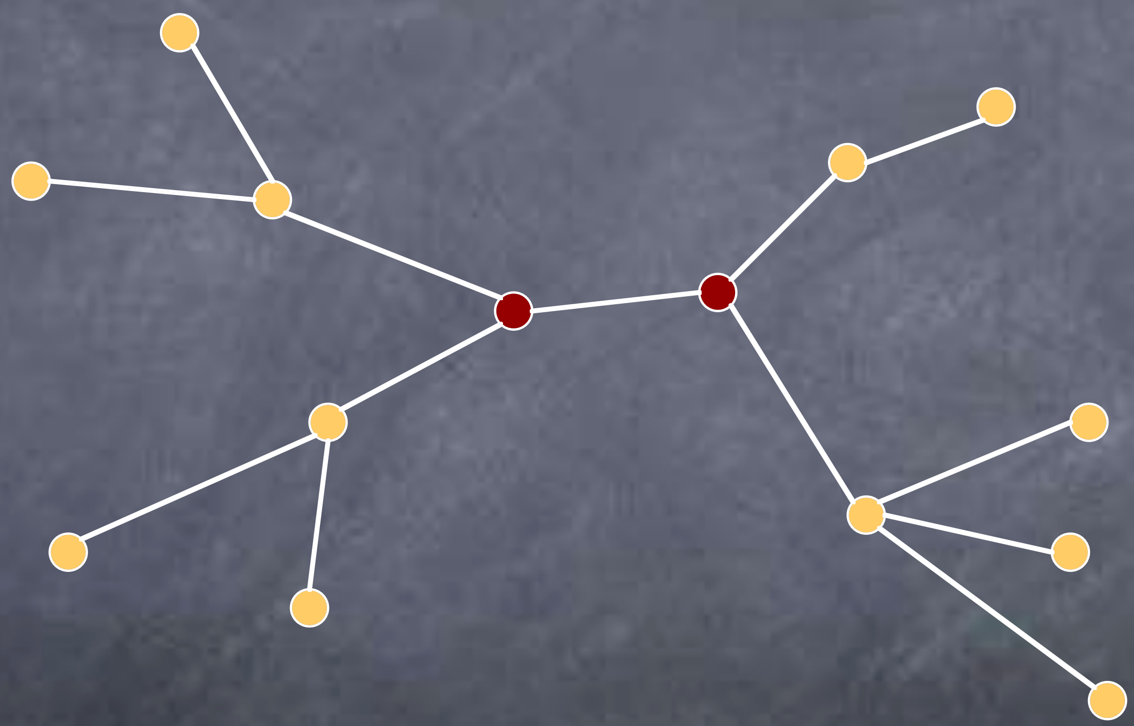
Limitations of this approach

E.g.: IEEE 1394 Leader election protocol



Limitations of this approach

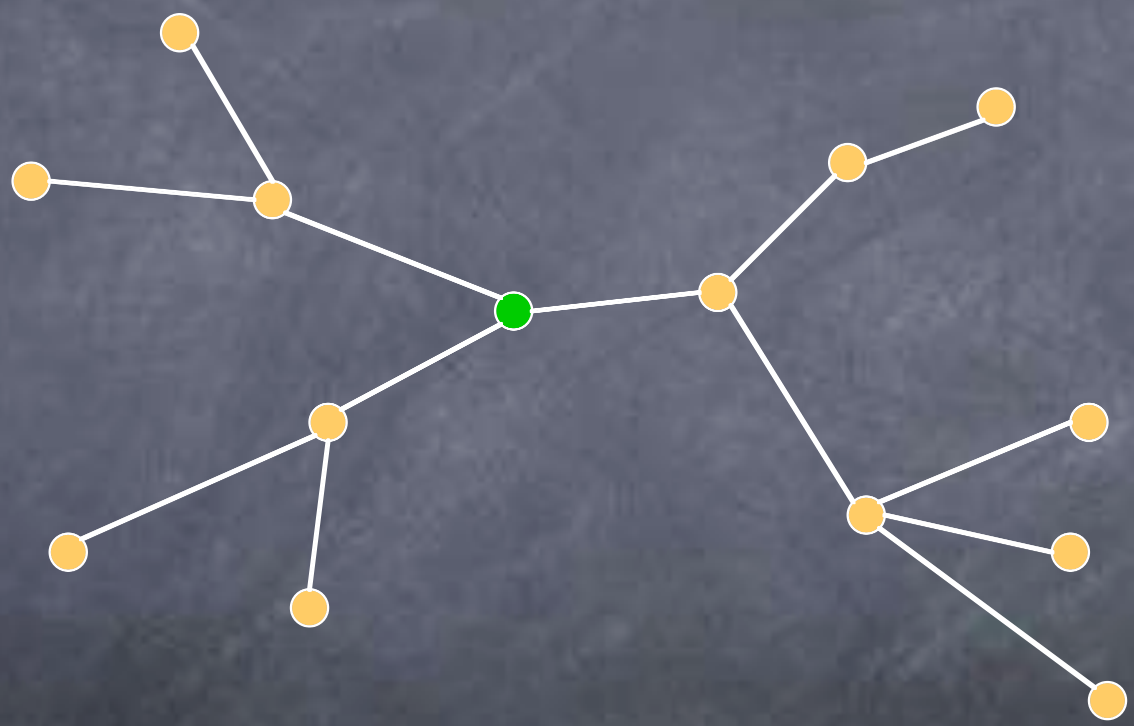
E.g.: IEEE 1394 Leader election protocol



Root contention!

Limitations of this approach

E.g.: IEEE 1394 Leader election protocol



It is solved by “flipping coins”

Limitations of this approach

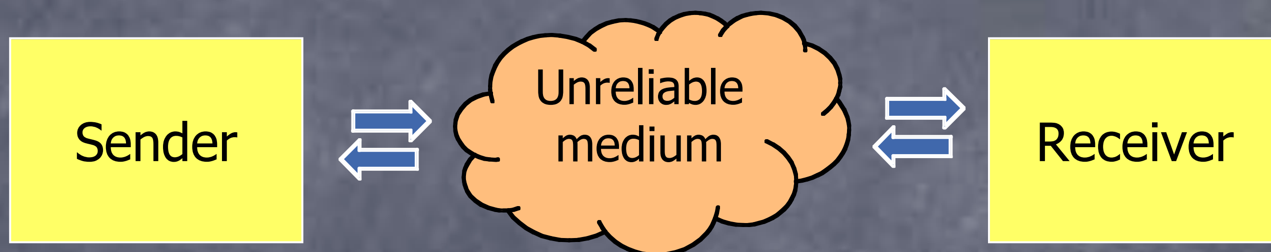
- Many times, correction cannot be established in a usual bivalued (modal) logic.
- Nevertheless, the validity of a property can be quantified through a probability value.
- E.g.
  - Bounded Retransmission Protocol en Philips RC6
  - Binary Exponential Backoff Algorithm en IEEE 802.3 "Ethernet"

CONICET

UNC

Limitations of this approach

Suppose that a file is transmitted using the ABP or a sliding window protocol



G (send(msg) => F rcv(msg))

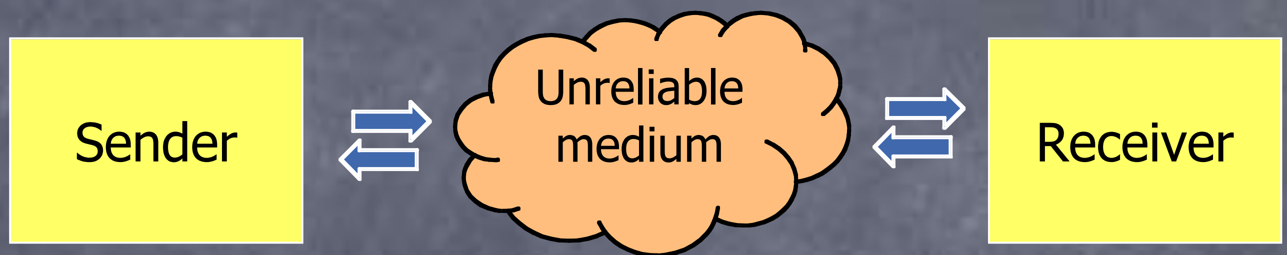
but this is under the assumption that an infinite number of retrials is allowed

CONICET

UNC

Limitations of this approach

Suppose that a file is transmitted using the ABP or a sliding window protocol



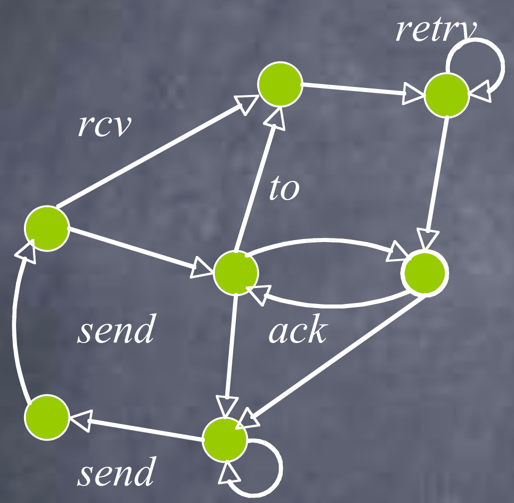
G (send(msg) => F rcv(msg))

What if only a bounded number of retransmissions is allowed? (e.g. BRP)

CONICET

UNC

Limitations of this approach



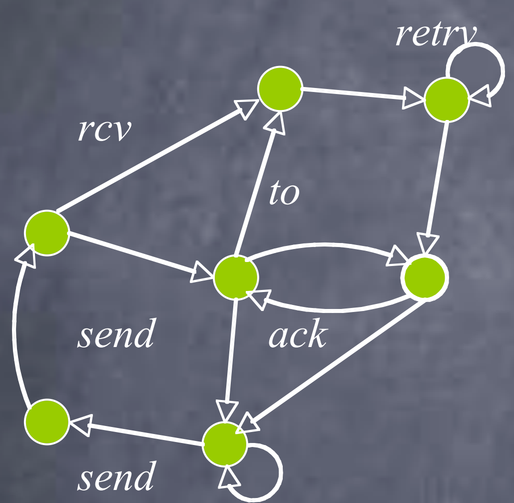
Properties are either true or false

$\models$ G (send(msg) => F rcv(msg))

Non-deterministic behavior

Limitations of this approach

The truth value should be probabilistically quantified

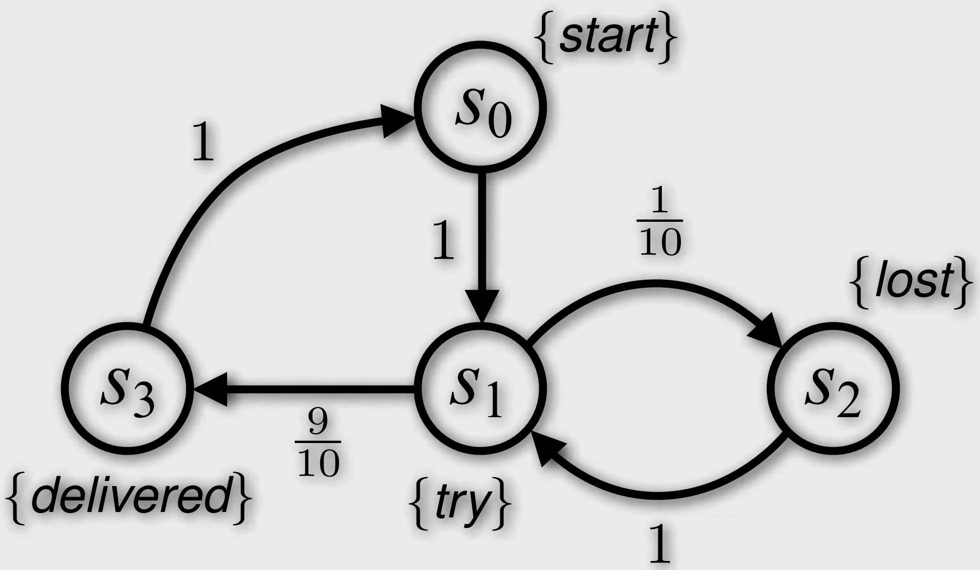


$\models$ G (send(msg) => F rcv(msg))

Non-deterministic behavior

Probabilistic behavior should also be considered

Fully probabilistic systems (Markov Chain)



$(S, \mathbf{P}, s_0, L)$

set of states with initial state s_0

$\mathbf{P} : S \times S \to [0,1]$
is the probabilistic transition function, s.t. $\forall s \in S$, $\sum_{s' \in S} \mathbf{P}(s, s') = 1$, and

$L : S \to \mathscr{P}(AP)$ labelling function, where AP is the set of atomic propositions.

$S = \{s_0, s_1, s_2, s_3\}$

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{10} & \frac{9}{10} \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$L(s_0) = \{start\}$
$L(s_1) = \{try\}$
$L(s_2) = \{lost\}$
$L(s_3) = \{delivered\}$

Probability of a property

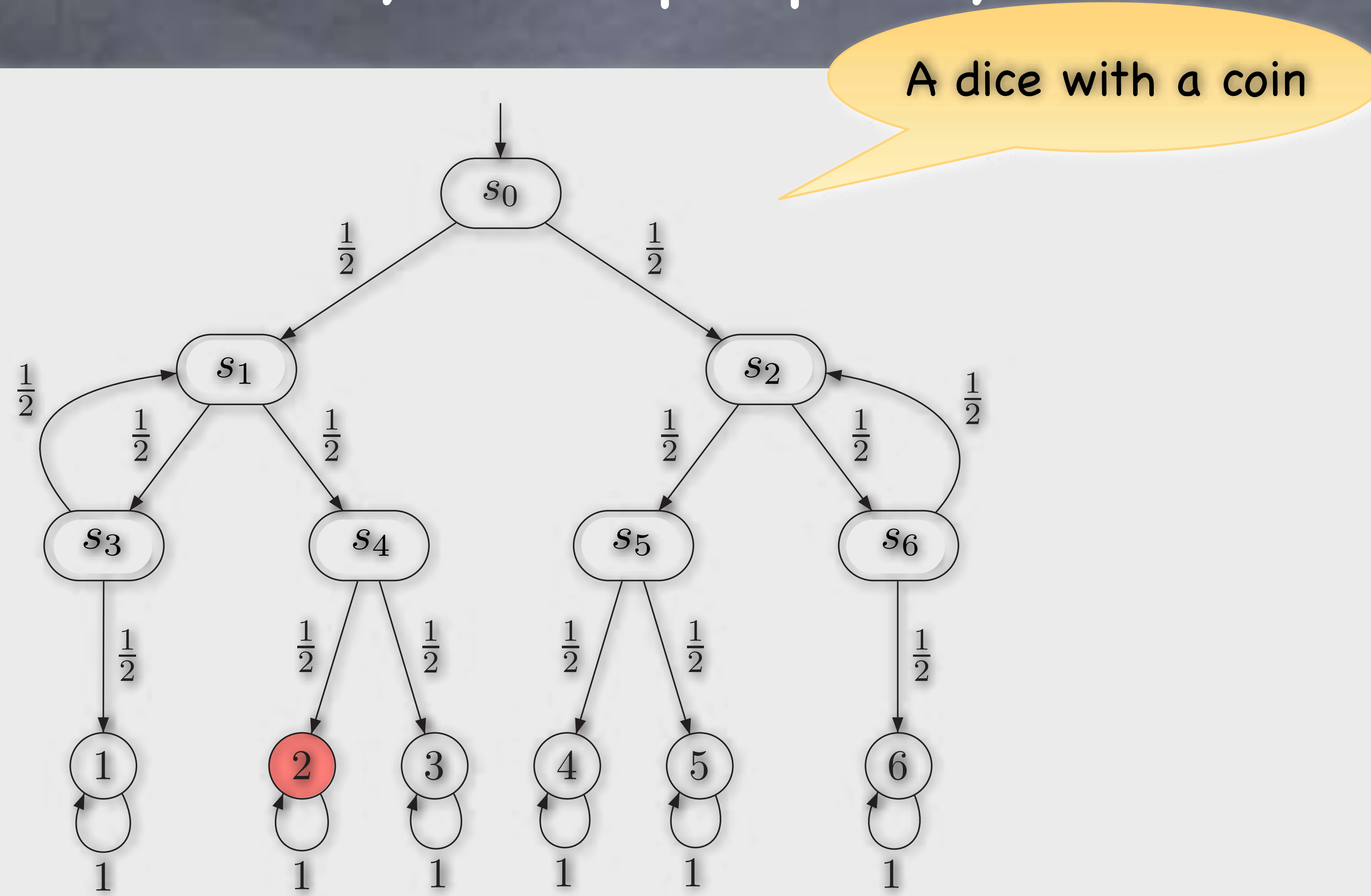
- Models contain probabilistic information (e.g. a decision made by tossing a coin, the probability of loosing a message).
- The validity of a temporal fomula (e.g. LTL) is quantified with a probability value in [0,1] (instead of a boolean).

Prob(F ●) =

$$= 0.5*0.4 + 0.5*0.2 + 0.5*0.7$$

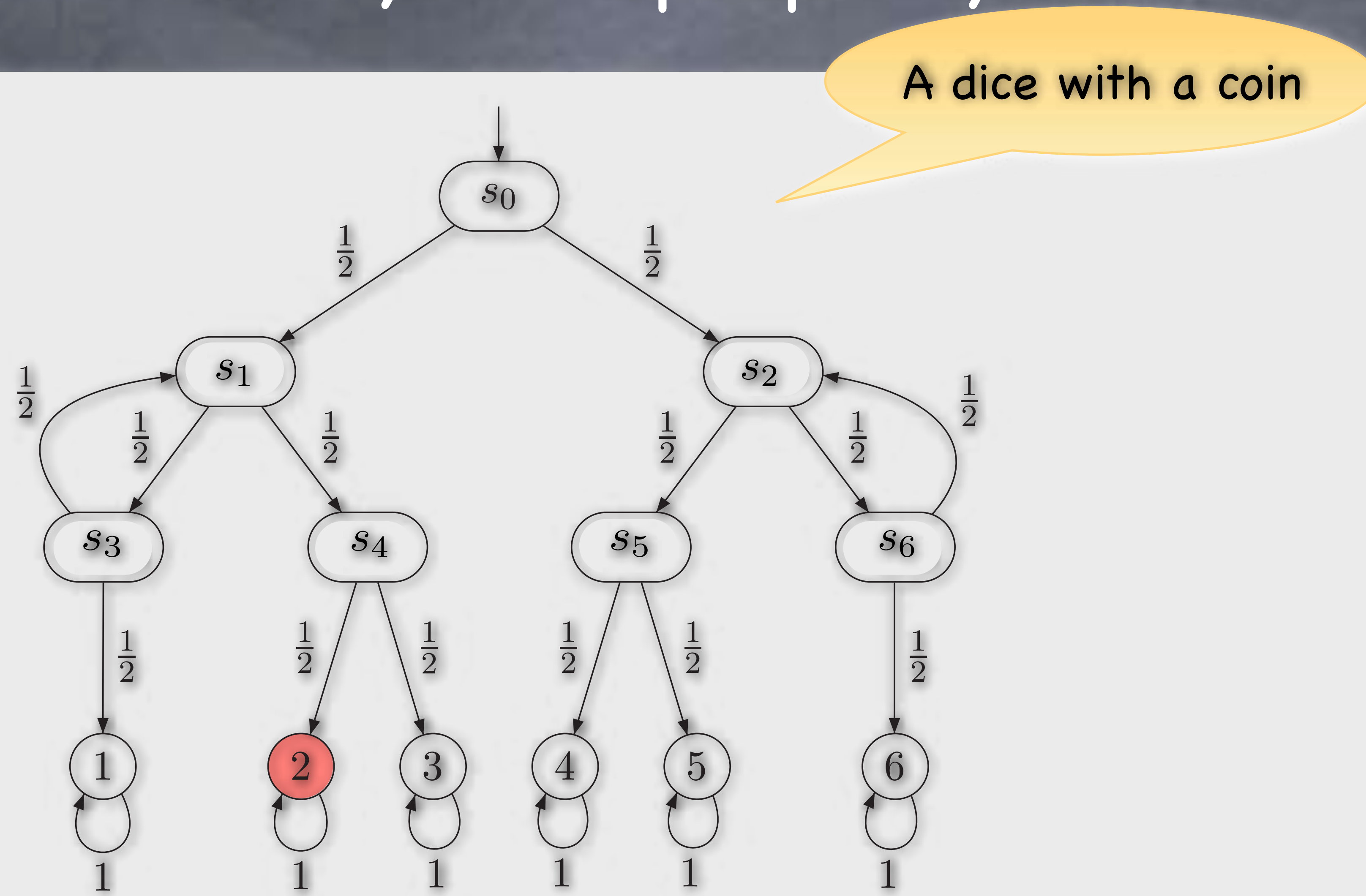
$$= 0.65$$

Probability of a property



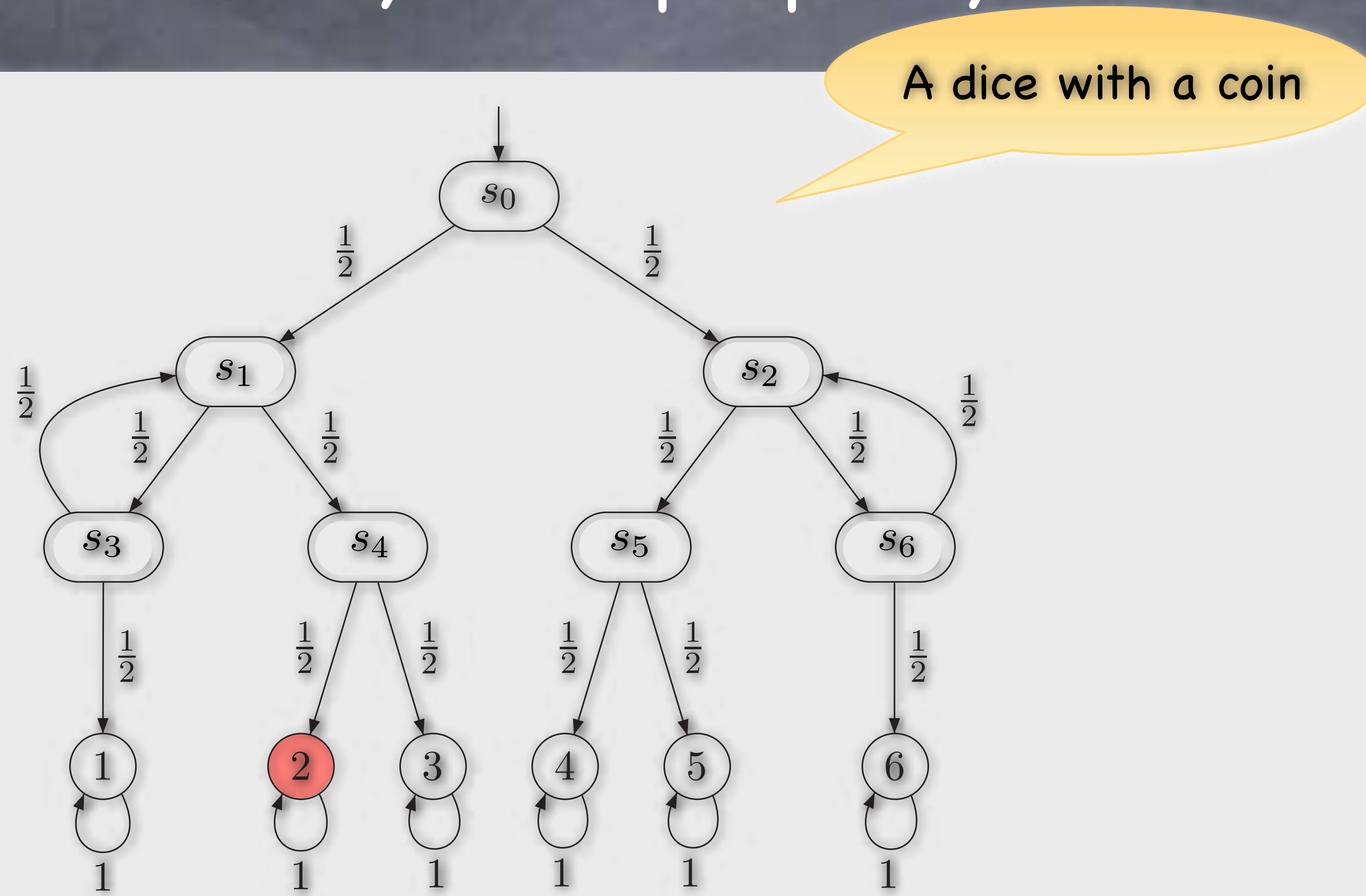
¿$P($ F 2 $)$?

Probability of a property



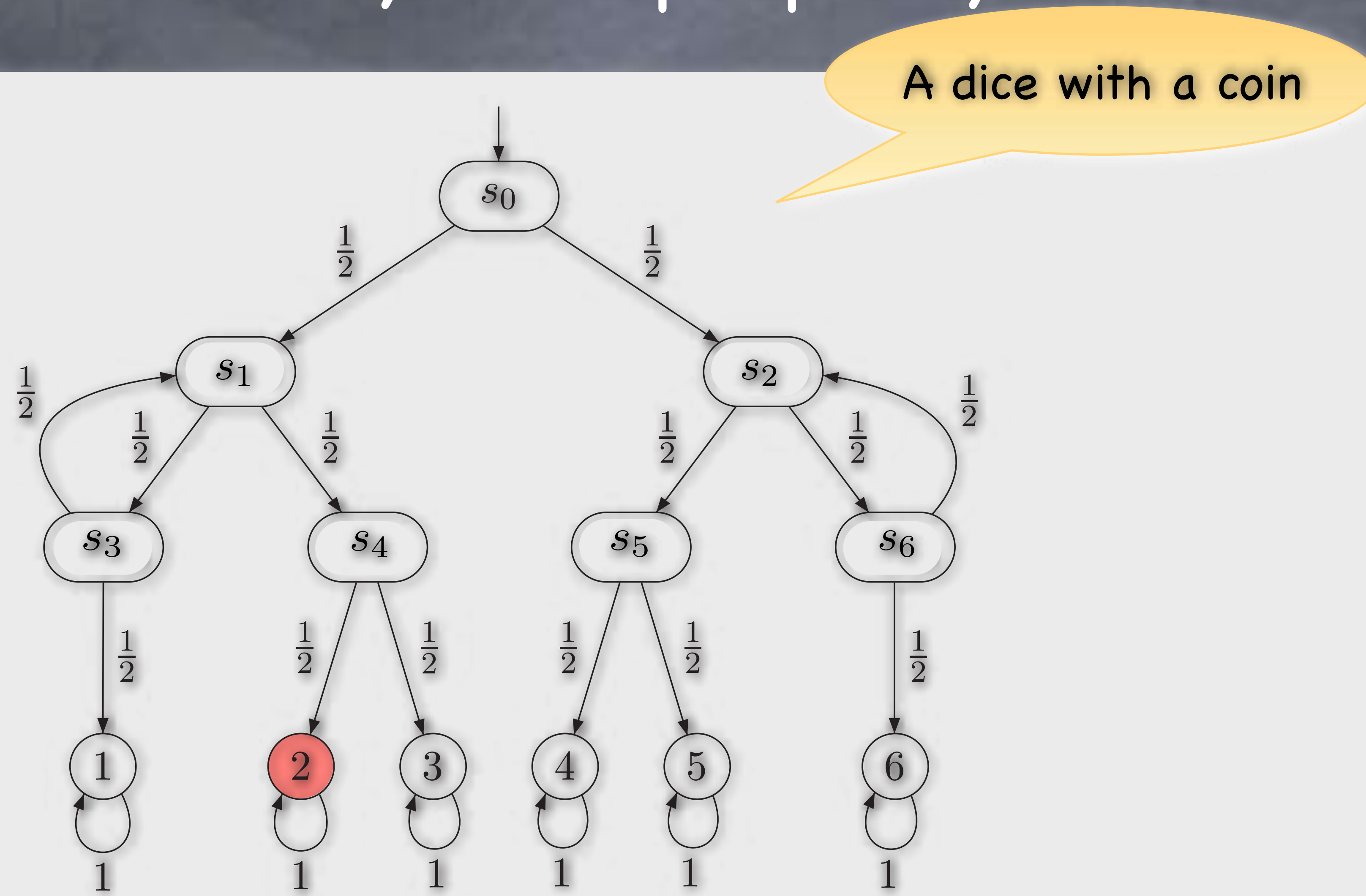
$$\underbrace{P(s_0 s_1 s_4 2)}_{\mathbf{P}(s_0, s_1) \cdot \mathbf{P}(s_1, s_4) \cdot \mathbf{P}(s_4, 2)} + P(s_0 s_1 s_3 s_1 s_4 2) + P(s_0 s_1 s_3 s_1 s_3 s_1 s_4 2) + P(s_0 s_1 s_3 s_1 s_3 s_1 s_3 s_1 s_4 2) + \cdots$$

Probability of a property



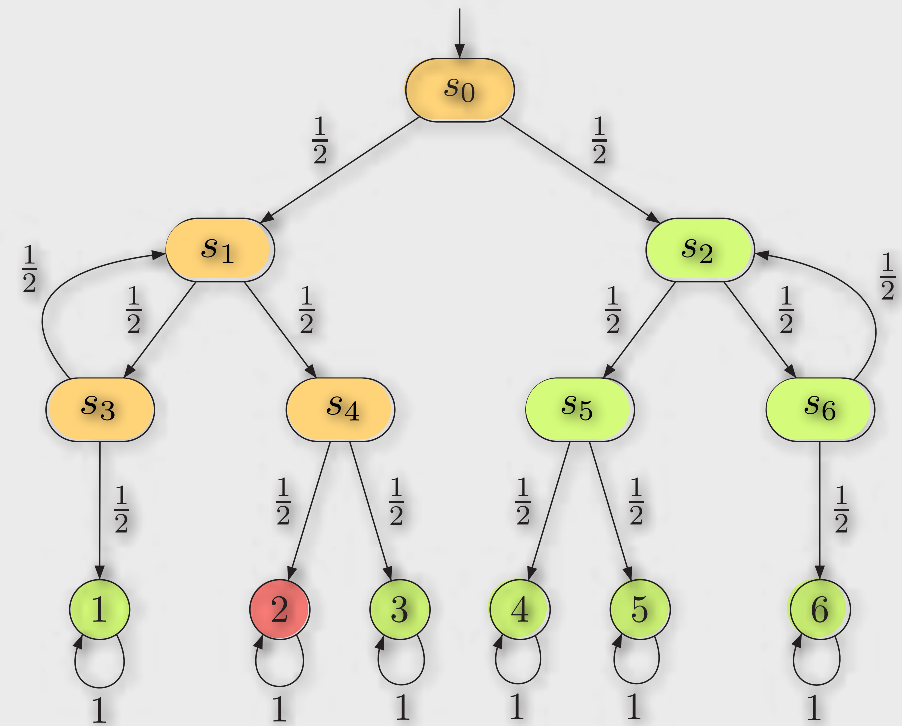
$$\underbrace{P(s_0 s_1 s_4 2)}_{\frac{1}{8}} + \underbrace{P(s_0 s_1 s_3 s_1 s_4 2)}_{\frac{1}{32}} + \underbrace{P(s_0 s_1 s_3 s_1 s_3 s_1 s_4 2)}_{\frac{1}{128}} + \underbrace{P(s_0 s_1 s_3 s_1 s_3 s_1 s_3 s_1 s_4 2)}_{\frac{1}{512}} + \cdots$$

Probability of a property



$$P_{s_0}(\,\mathsf{F}\,2\,) \;=\; \sum_{n>0} \mathbf{P}(s_0 s_1 (s_3 s_1)^n s_4 2) \;=\; \sum_{n>0} \frac{1}{2^{2n+1}} \;=\; \frac{1}{6}$$

Probabilistic Model Checking in fully probabilistic models



$$P_{s_2}(\mathsf{F}\ 2) = P_{s_5}(\mathsf{F}\ 2) = P_{s_6}(\mathsf{F}\ 2) = 0$$

$$P_1(\mathsf{F}\ 2) = P_3(\mathsf{F}\ 2) = P_4(\mathsf{F}\ 2) = 0$$

$$P_5(\mathsf{F}\ 2) = P_6(\mathsf{F}\ 2) = 0$$

$$P_2(\mathsf{F}\ 2) = 1$$

$$P_{s_0}(\mathsf{F}\ 2) = \frac{1}{2}\, P_{s_1}(\mathsf{F}\ 2) + \frac{1}{2}\, P_{s_2}(\mathsf{F}\ 2)$$

$$P_{s_1}(\mathsf{F}\ 2) = \frac{1}{2}\, P_{s_3}(\mathsf{F}\ 2) + \frac{1}{2}\, P_{s_4}(\mathsf{F}\ 2)$$

$$P_{s_3}(\mathsf{F}\ 2) = \frac{1}{2}\, P_{s_1}(\mathsf{F}\ 2) + \frac{1}{2}\, P_1(\mathsf{F}\ 2)$$

$$P_{s_4}(\mathsf{F}\ 2) = \frac{1}{2} P_2\, (\mathsf{F}\ 2) + \frac{1}{2} P_3\, (\mathsf{F}\ 2)$$

Using DFS, we can calculate whether 2 is reachable with probability 0

Probabilistic Model Checking in fully probabilistic models

In general:

$$x_s = \sum_{t \in S} \mathbf{P}(s,t) \cdot x_t \qquad \text{if } s \in Pr^{>0}(B) \setminus B$$

$$x_s = 1 \qquad \text{if } s \in B$$

$$x_s = 0 \qquad \text{if } s \notin Pr^{>0}(B)$$

B is the set of goal states

The set of states that reach B with some probability

It is solved with standard numeric techniques (Jacobi, Gauss-Seidel)

CONICET

UNC

The need of non-determinism

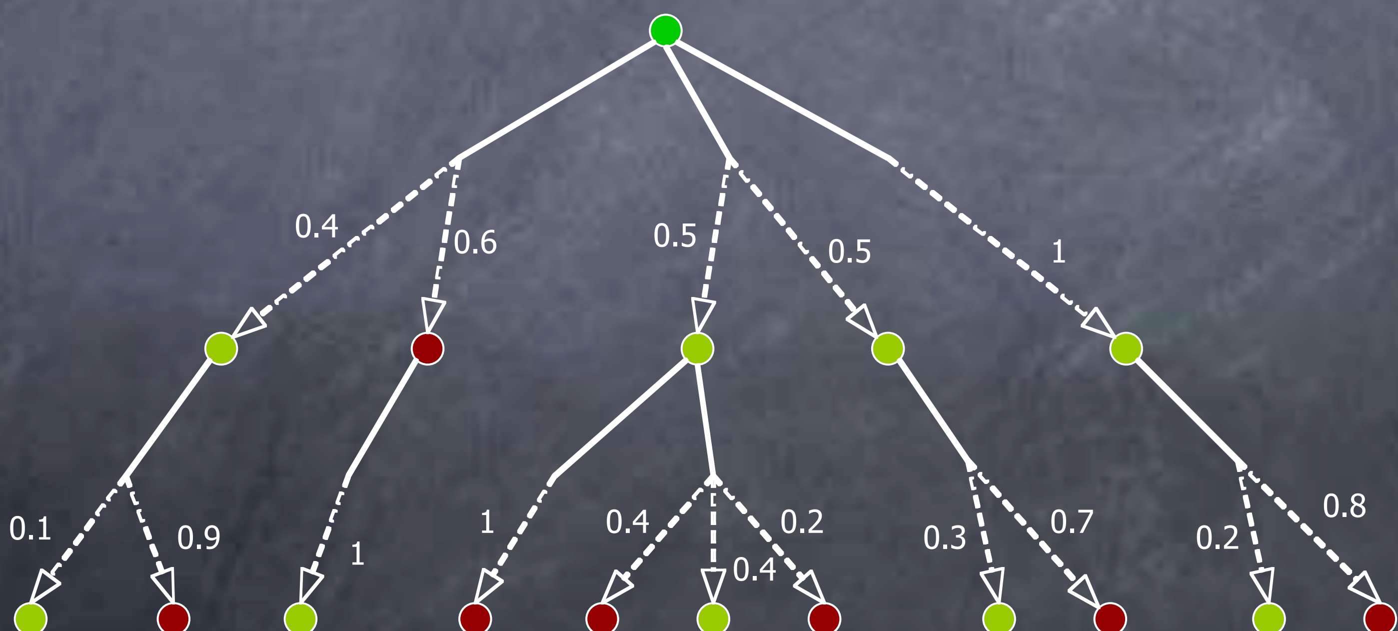
- Parallel composition / Distributed components
  - probabilities within a single component are easy to estimate,
  - relative probabilities of events located geographically distant depend on a highly unpredictable global state.
- Underspecification
  - some probabilities are unknown at early stage of modeling.
- Abstraction
  - models are abstract representations of the system under study.
- Control synthesis
  - intentional underspecification to synthesize optimal decisions.

CONICET

UNC

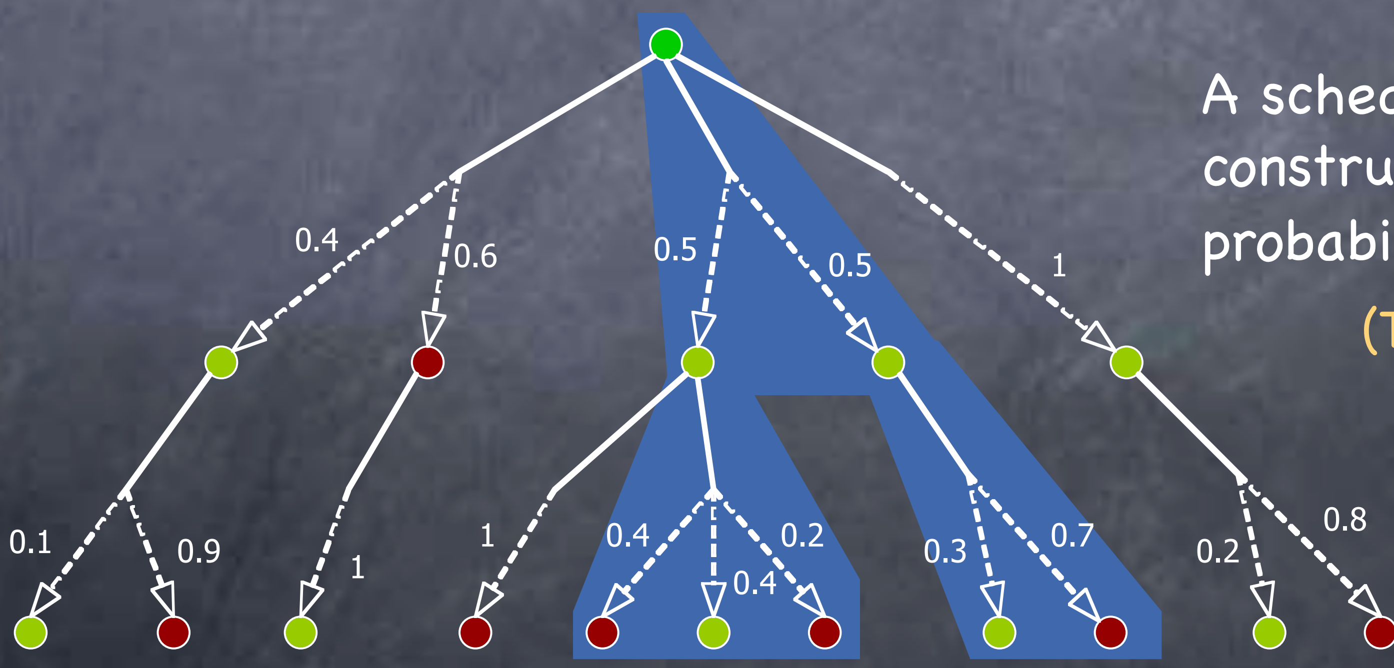
Probability of a property

- To calculate probabilities in this setting, non-determinism has to be resolved.
- Schedulers are functions that select the next transition according to the past execution.



Probability of a property

- To calculate probabilities in this setting, non-determinism has to be resolved.
- Schedulers are functions that select the next transition according to the past execution.



A scheduler constructs a fully probabilistic tree

(There are also randomized variants)

CONICET

UNC

Probability of a property

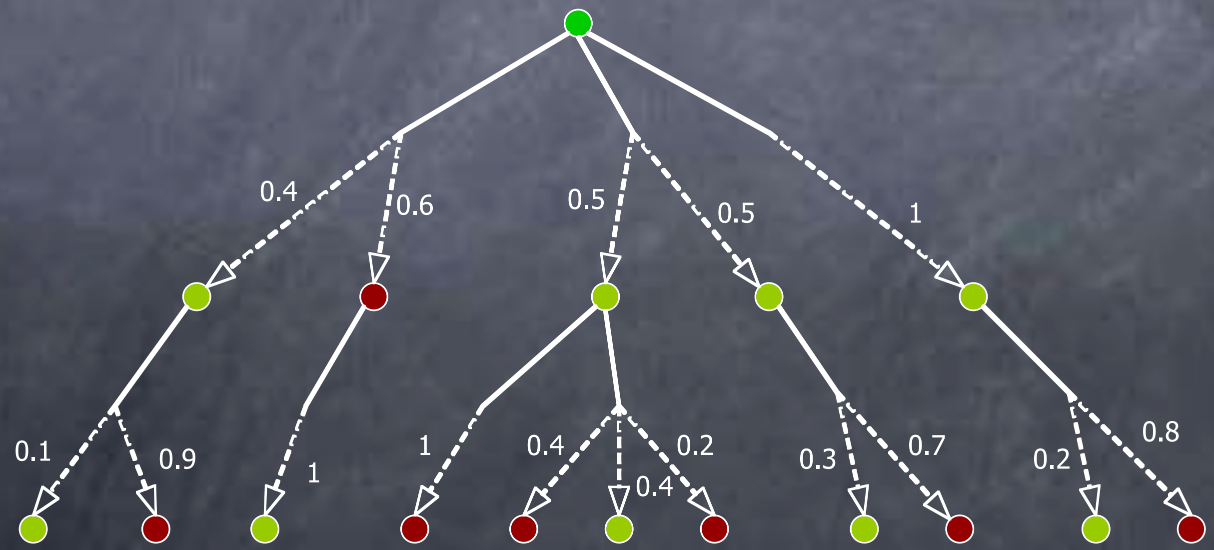
An LTL formula has associated two values:

- The maximum probability under all schedulers

$$P_{max}(F \bullet) = 0.96$$

- The minimun probability under all schedulers

$$P_{min}(F \bullet) = 0.65$$



CONICET

UNC

Probability of a property

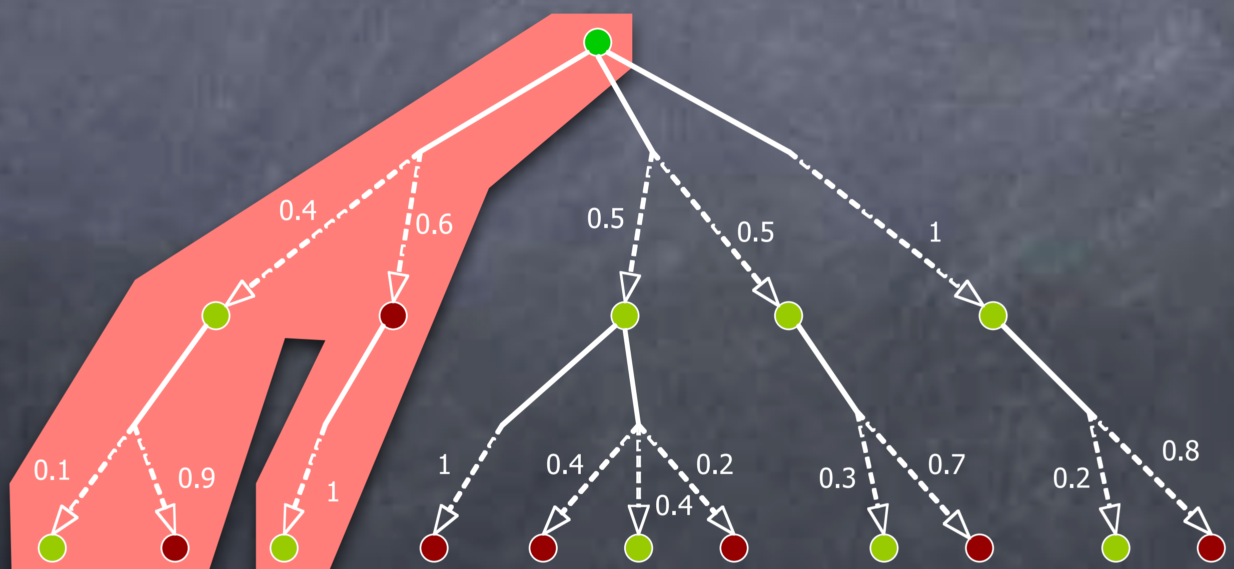
An LTL formula has associated two values:

- The maximum probability under all schedulers

  $P_{max}(\ F\ \bullet\ ) = 0.96$

- The minimun probability under all schedulers

  $P_{min}(\ F\ \bullet\ ) = 0.65$



Probability of a property

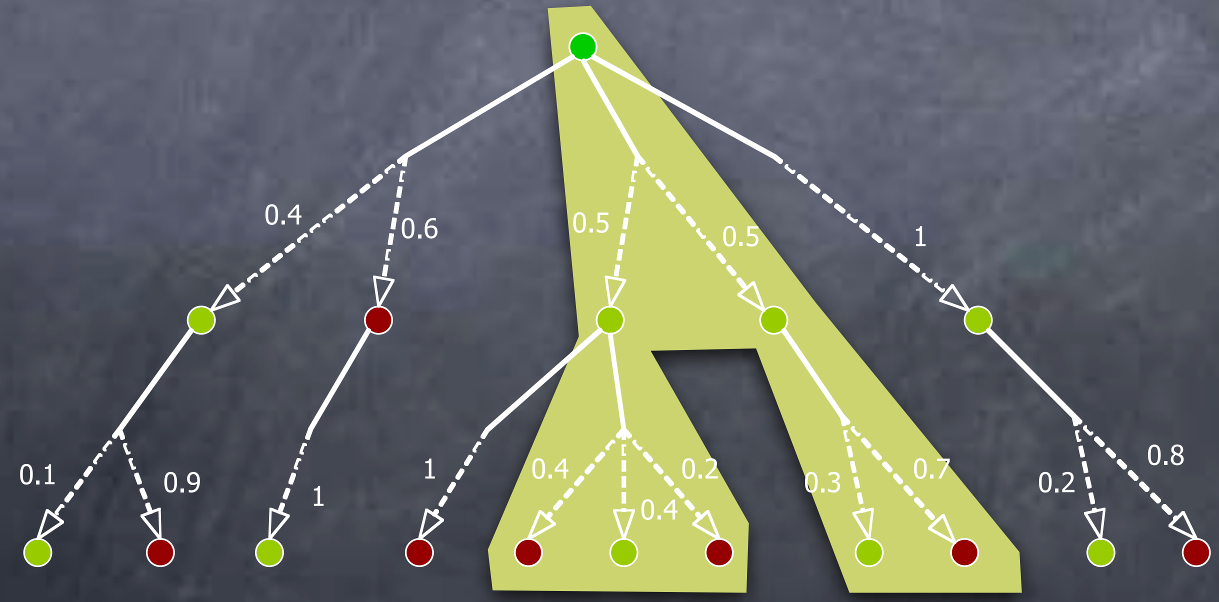
An LTL formula has associated two values:

- The maximum probability under all schedulers

  $P_{max}(F \bullet) = 0.96$

- The minimun probability under all schedulers

  $P_{min}(F \bullet) = 0.65$



CONICET

UNC

Probability of

Randomized and deterministic schedulers are equally expressive for max/min prob. of reach. properties

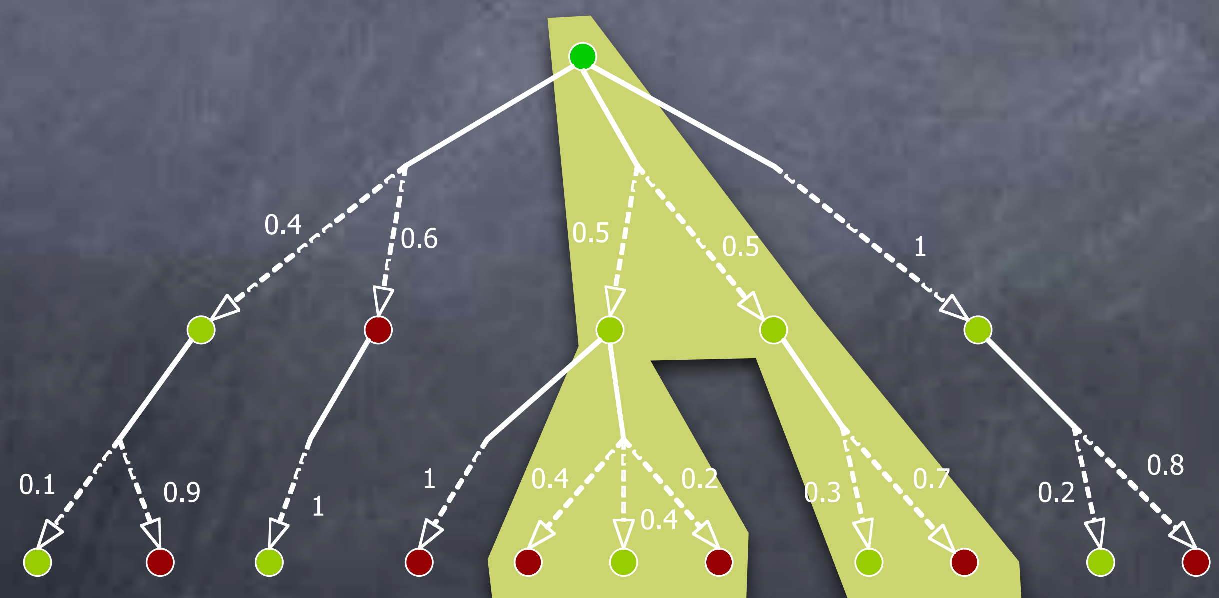
An LTL formula has associated tw values:

- The maximum probability under all schedulers

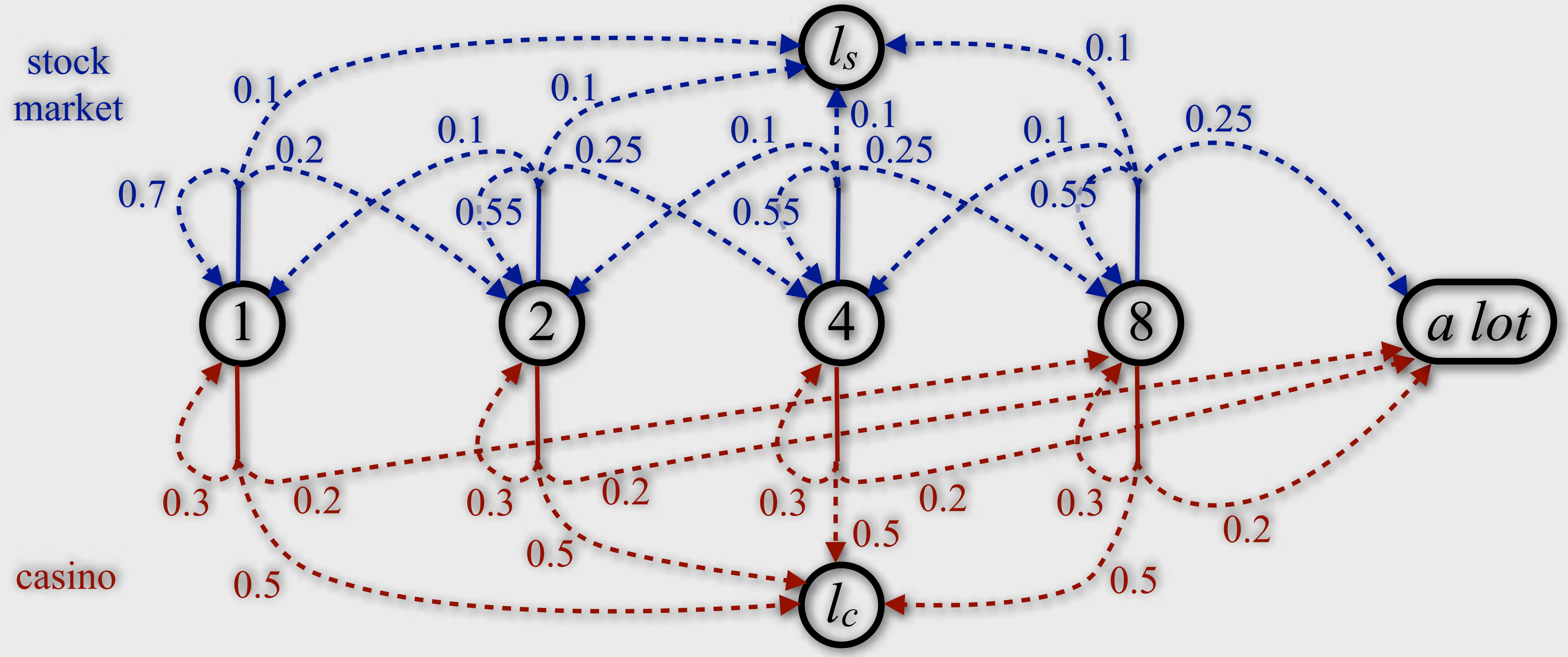
$$P_{max}(F \bullet) = 0.96$$

- The minimun probability under all schedulers

$$P_{min}(F \bullet) = 0.65$$



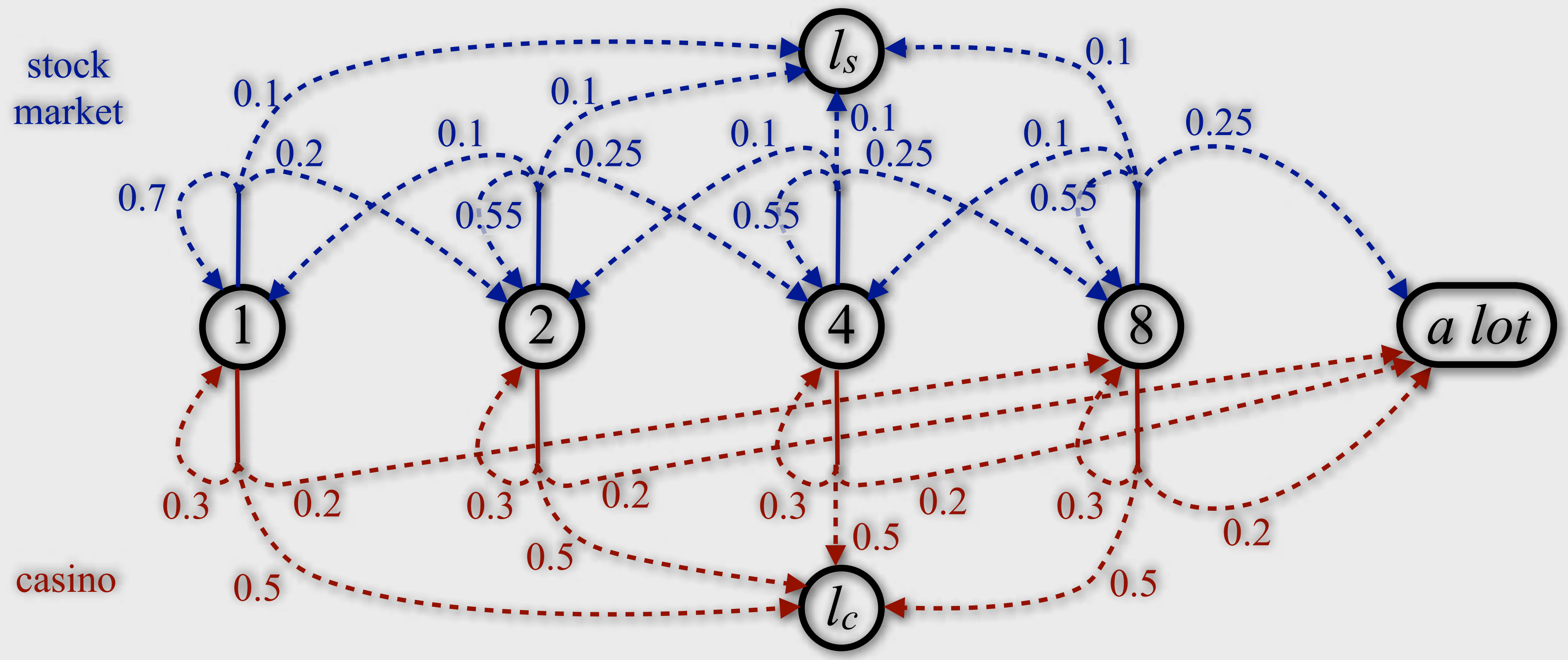
Markov decision processes



$$(S, \{\mathbf{P}_a\}_{a \in A}, s_0, L)$$

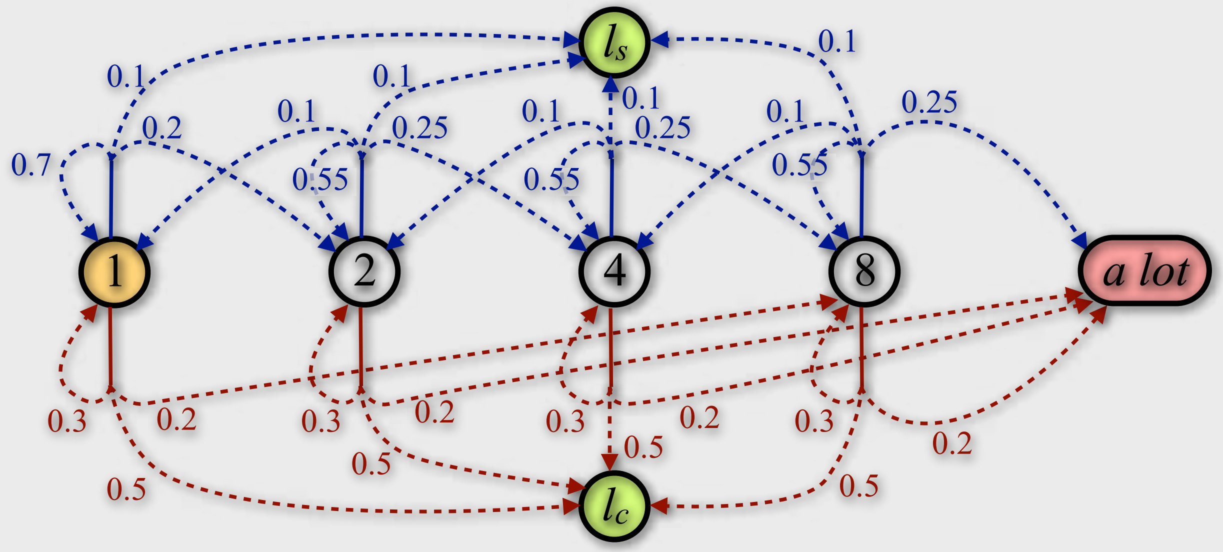
The structure is as before, only that we have a family of matrices, one for each posible decision

Markov decision processes



What is the maximum probability of obtaining the desired amount of money?

Model checking
Markov decision processes

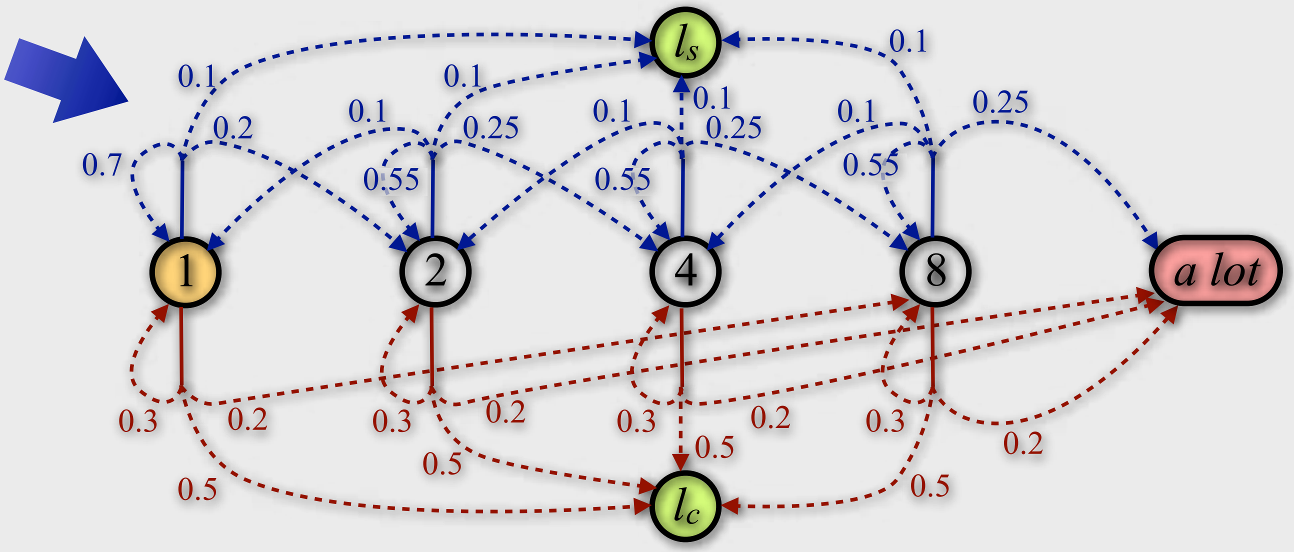


P_s^+ is a shorthand for $P_s^{max}(\mathsf{F}\ al)$

$$P_{l_s}^+ = P_{l_c}^+ = 0$$

$$P_{al}^+ = 1$$

Model checking
Markov decision processes



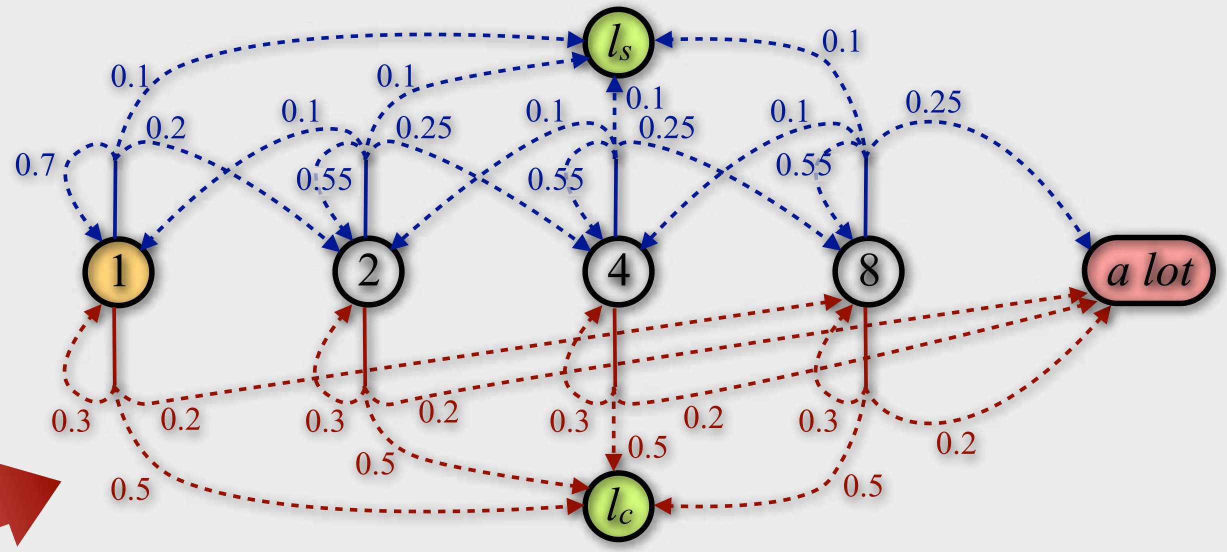
P_s^+ is a shorthand for $P_s^{max}(\mathsf{F}\ al)$

$P_{l_s}^+ = P_{l_c}^+ = 0$

$P_{al}^+ = 1$

$P_1^+ = \quad 0.7P_1^+ + 0.2P_2^+ + 0.1P_{l_s}^+$

Model checking
Markov decision processes



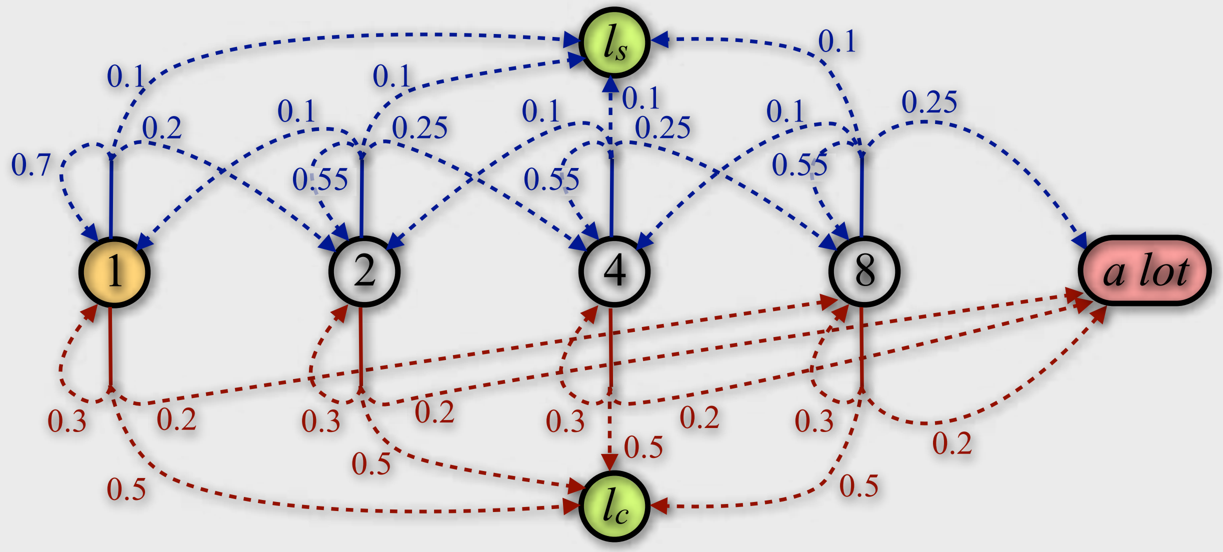
P_s^+ is a shorthand for $P_s^{max}(\mathsf{F}\ al)$

$P^+_{l_s} = P^+_{l_c} = 0$

$P^+_{al} = 1$

$P^+_1 = \quad 0.3P^+_1 + 0.2P^+_8 + 0.5P^+_{l_c}$

Model checking
Markov decision processes



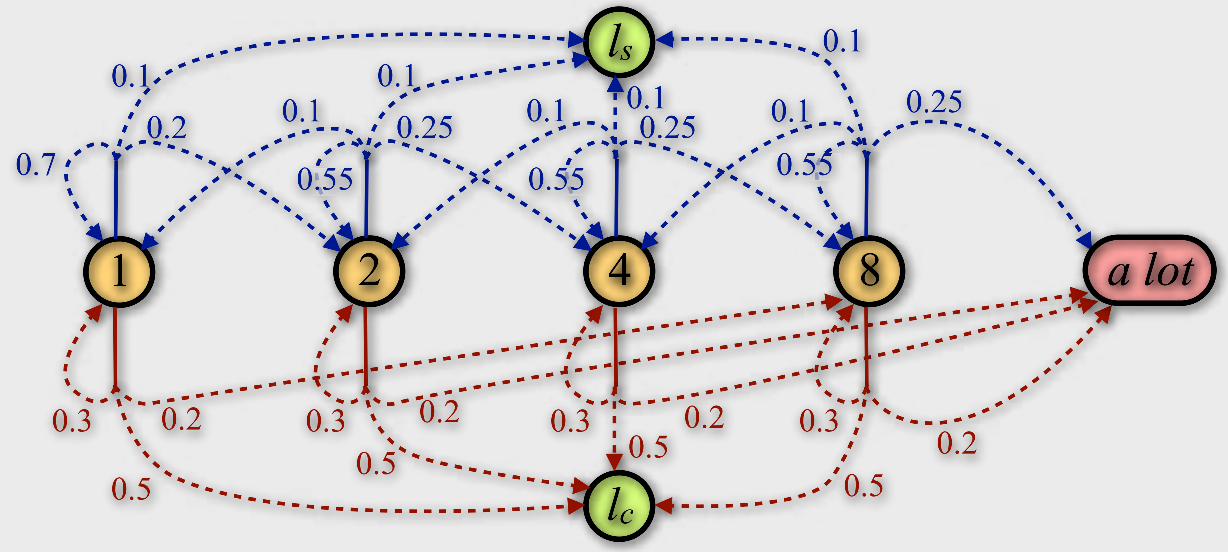
P_s^+ is a shorthand for $P_s^{max}(\mathsf{F}\ al)$

$$P_{l_s}^+ = P_{l_c}^+ = 0$$

$$P_{al}^+ = 1$$

$$P_1^+ = \max\left(\, 0.7P_1^+ + 0.2P_2^+ + 0.1P_{l_s}^+,\ 0.3P_1^+ + 0.2P_8^+ + 0.5P_{l_c}^+ \,\right)$$

Model checking
Markov decision processes



P_s^+ is a shorthand for $P_s^{max}(\mathsf{F}\ al)$

$$P_{l_s}^+ = P_{l_c}^+ = 0$$

$$P_{al}^+ = 1$$

$$P_1^+ = \max\left(0.7P_1^+ + 0.2P_2^+ + 0.1P_{l_s}^+,\ 0.3P_1^+ + 0.2P_8^+ + 0.5P_{l_c}^+ \right)$$

$$P_2^+ = \max\left(0.55P_2^+ + 0.25P_4^+ + 0.1P_1^+ + 0.1P_{l_s}^+,\ 0.3P_2^+ + 0.2P_{al}^+ + 0.5P_{l_c}^+ \right)$$

$$P_4^+ = \max\left(0.55P_4^+ + 0.25P_8^+ + 0.1P_2^+ + 0.1P_{l_s}^+,\ 0.3P_4^+ + 0.2P_{al}^+ + 0.5P_{l_c}^+ \right)$$

$$P_8^+ = \max\left(0.55P_8^+ + 0.25P_{al}^+ + 0.1P_4^+ + 0.1P_{l_s}^+,\ 0.3P_8^+ + 0.2P_{al}^+ + 0.5P_{l_c}^+ \right)$$

Model checking
Markov decision processes

In general:

$$x_s = \max_{a \in A} \sum_{t \in S} \mathbf{P}_a(s,t) \cdot x_t \qquad \text{if } s \in Pr^{>0}(B) \setminus B$$

$$x_s = 1 \qquad \text{if } s \in B$$

$$x_s = 0 \qquad \text{if } s \notin Pr^{>0}(B)$$

B is the set of goal states

The set of states that may reach B with some probability

Linear optimization problem.
Solved with standard numerical analysis techniques

CONICET

UNC

LTL reduced to reachability

LTL = propositional logic + temporal modalities:

- G φ : "φ holds globally"
- F φ : "Finally φ holds"
- φ U ψ : "φ holds until ψ holds"

E.g.:

$$\mathsf{G}\ (\ \textit{send-msg} \Rightarrow \mathsf{F}\ \textit{rcv-msg}\ )$$

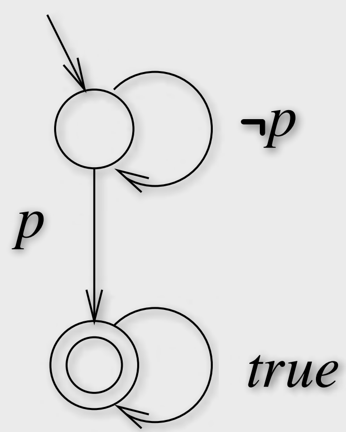
CONICET

UNC

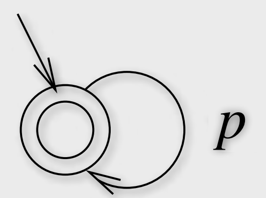
LTL reduced to reachability

Every LTL formula can be translate to a Büchi Automaton that represents the accepting behaviour.

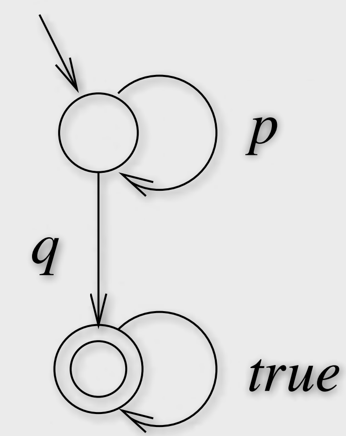
$\mathsf{F}\, p$



$\mathsf{G}\, p$



$p \,\mathsf{U}\, q$



CONICET

UNC

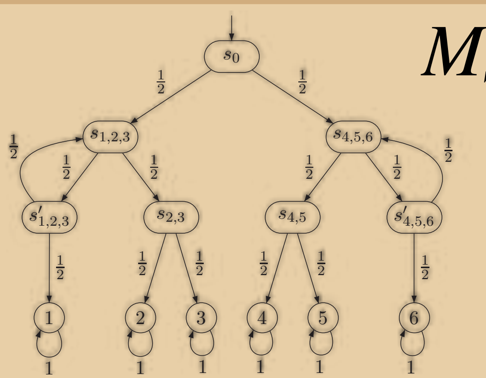
$P_S(\phi) = ?$

S

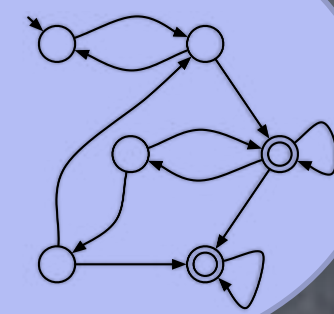
```
dtmc
module die
  s : [0..7] init 0;
  d : [0..6] init 0;
  [] s=0 -> 0.5 : (s'=1) + 0.5 : (s'=2);
  [] s=1 -> 0.5 : (s'=3) + 0.5 : (s'=4);
  [] s=2 -> 0.5 : (s'=5) + 0.5 : (s'=6);
  [] s=3 -> 0.5 : (s'=1) + 0.5 : (s'=7) & (d'=1);
  [] s=4 -> 0.5 : (s'=7) & (d'=2) + 0.5 : (s'=7) & (d'=3);
  [] s=5 -> 0.5 : (s'=7) & (d'=4) + 0.5 : (s'=7) & (d'=5);
  [] s=6 -> 0.5 : (s'=2) + 0.5 : (s'=7) & (d'=6);
  [] s=7 -> (s'=7);
endmodule
```

ϕ : $\Box \Diamond crit_1 \wedge \Box \Diamond crit_2$

M_S



A_ϕ



Compose M_S with A_ϕ

Calculate probability of reaching accepting BSCCs in $M_S \times A_\phi$

CONICET

UNC

$P_S(\phi) = ?$

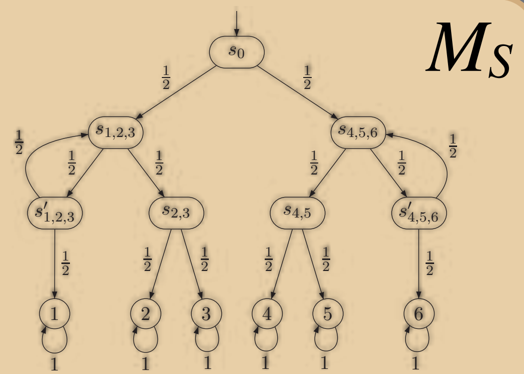
```
dtmc
module die
  s : [0..7] init 0;
  d : [0..6] init 0;
  [] s=0 -> 0.5 : (s'=1) + 0.5 : (s'=2);
  [] s=1 -> 0.5 : (s'=3) + 0.5 : (s'=4);
  [] s=2 -> 0.5 : (s'=5) + 0.5 : (s'=6);
  [] s=3 -> 0.5 : (s'=1) + 0.5 : (s'=7) & (d'=1);
  [] s=4 -> 0.5 : (s'=7) & (d'=2) + 0.5 : (s'=7) & (d'=3);
  [] s=5 -> 0.5 : (s'=7) & (d'=4) + 0.5 : (s'=7) & (d'=5);
  [] s=6 -> 0.5 : (s'=2) + 0.5 : (s'=7) & (d'=6);
  [] s=7 -> (s'=7);
endmodule
```

S

$\phi : \quad \Box \Diamond crit_1 \wedge \Box \Diamond crit_2$

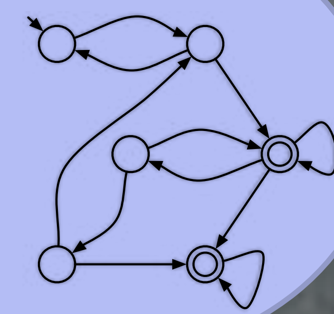
Modelling

M_S



Automatic

A_ϕ



Compose M_S with A_ϕ

Calculate probability of reaching accepting BSCCs in $M_S \times A_\phi$

CONICET

UNC

Highlights on Fundamentals of Probabilistic Model Checking

- Vardi '85
  - Qualitative MC on deterministic and non-deterministic PTSs
- Courcoubetis & Yanakakis '88
  - Quantitative MC on non-deterministic PTSs using LTL and lower/upper bounds
- Hansson & Jonsson '90
  - Quantitative MC on deterministic PTSs introducing PCTL
- Bianco & de Alfaro '95
  - Quantitative MC on non-deterministic PTSs using PCTL*
- de Alfaro, Kwiatkowska, Norman, Parker, & Segala '2000
  - Symbolic quantitative MC on non-deterministic PTSs

Highlights on Fundamentals of Probabilistic Model Che

1st. algorithm to qualitative MC MDPs

- Vardi '85
  - Qualitative MC on deterministic and non-determ

1st. algorithm for probabilistic MC

- Courcoubetis & Yanakakis '88
  - Quantitative MC on non-deterministic PTSs using LTL upper bounds

1st. modalities with probabilities

- Hansson & Jonsson '90
  - Quantitative MC on deterministic PTSs introducing PC

1st. "clever" algorithm

- Bianco & de Alfaro '95
  - Quantitative MC on non-deterministic PTSs using PCTL*
- de Alfaro, Kwiatkowska, Norman, Parker, & Segala '2000
  - Symbolic quantitative MC on non-deterministic PTSs

1st. efficient tool: PRISM

CONICET

… and more

- Model Checking Rewards properties

  [Andova, Hermanns & Katoen 2003]

- Model Checking CTMC & steady state properties

  [Baier, Havenkort, Hermanns & Katoen 2002]

- Model Checking CTMDP

  [Baier, Hermanns, Katoen & Havenkort 2004 / Baier, Hahn, Havenkort, Hermanns & Katoen 2013]

- Counterexample derivation

  [Aljazzar, Hermanns & Leue, 2005 / Han & Katoen 2007 / Andrés, D'Argenio, van Rossum 2008 / Damman, Han & Katoen 2008 / Jansen 2015]

CONICET

UNC

… and more

- Attacking the state explosion problem
  - Abstraction techniques

    [D'Argenio, Jeannet, Jensen, & Larsen, 2001 / Kwiatkowska, Norman, & Parker, 2006 / Wachter, Zhang, & Hermanns, 2007, 2008]
  - Partial order reduction

    [Baier, Ciesinski, & Größer, 2004 / D'Argenio & Niebert, 2004 / Baier, D'Argenio, & Größer, 2006 / Giro, D'Argenio, & Ferrer Fioriti, 2009]
- and much more:
  - Controller synthesis and games
  - Partial observation & distributed schedulers
  - Statistical Model Checking

CONICET

UNC

An Introduction to Probabilistic Model Checking

Pedro R. D'Argenio

Dependable Systems Group - FaMAF

Universidad Nacional de Córdoba
CONICET

http://dsg.famaf.unc.edu.ar

11º ERPEM, Rio Cuarto, Dec-2015

CONICET

UNC