On the Metric-based Approximate Minimization of Markov Chains*

Giovanni Bacci, **Giorgio Bacci**, Kim G. Larsen, Radu Mardare
Aalborg University

ICALP 2017

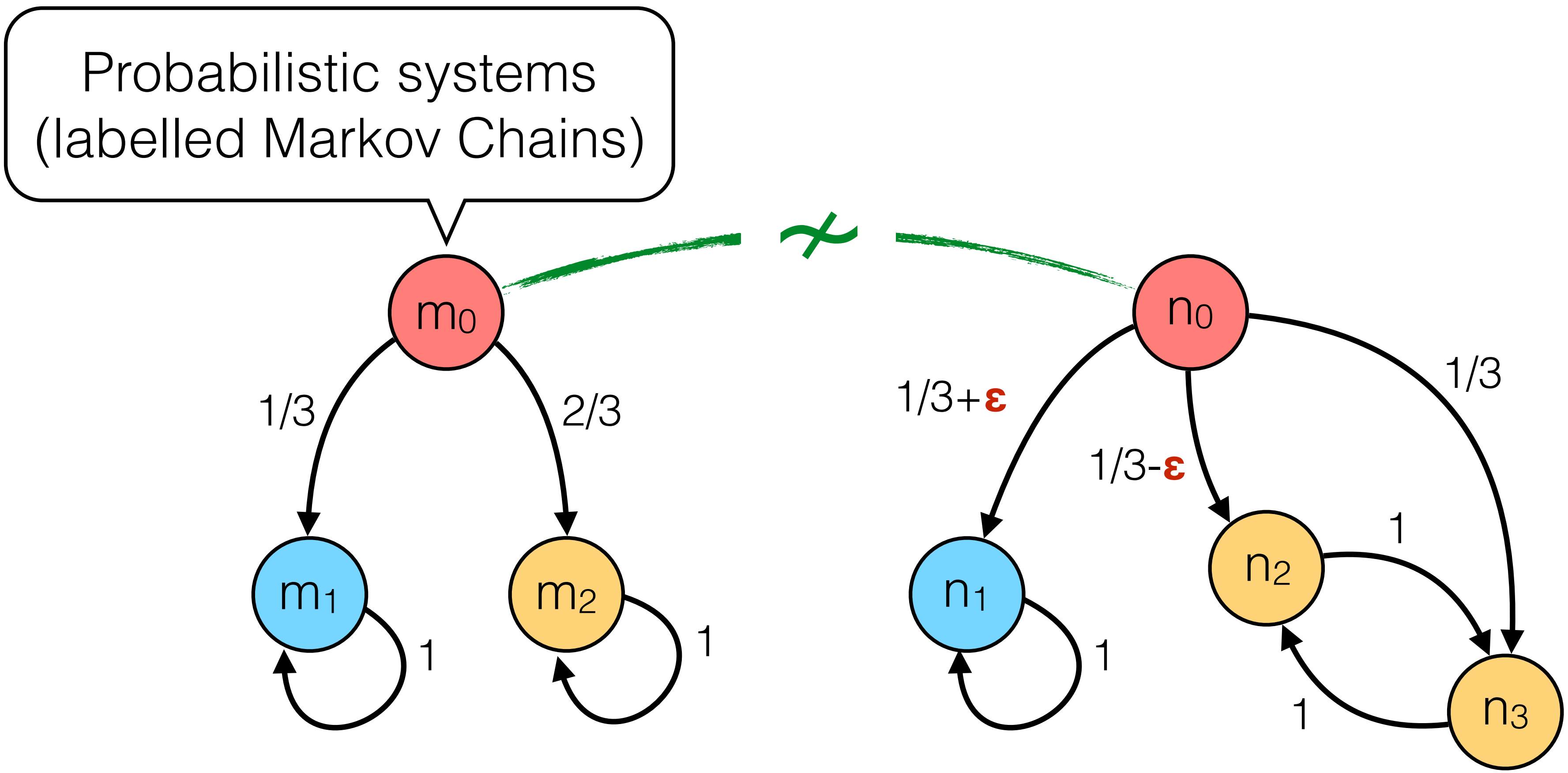
Warsaw, 11th July 2017

Introduction

- **Moore‘56, Hopcroft‘71:** Minimization algorithm for DFA (*partition refinement wrt Myhill-Nerode equiv.*)
- Minimization via partition refinement:
  - **Kanellakis-Smolka’83:** minimization of LTSs wrt Milner’s strong bisimulation
  - **Baier’96:** minimization of MCs wrt Larsen-Skou probabilistic bisimulation
  - **Alur et al.’92, Yannakakis-Lee’97:** minimization of timed & real-time transition systems.
  - and many more…

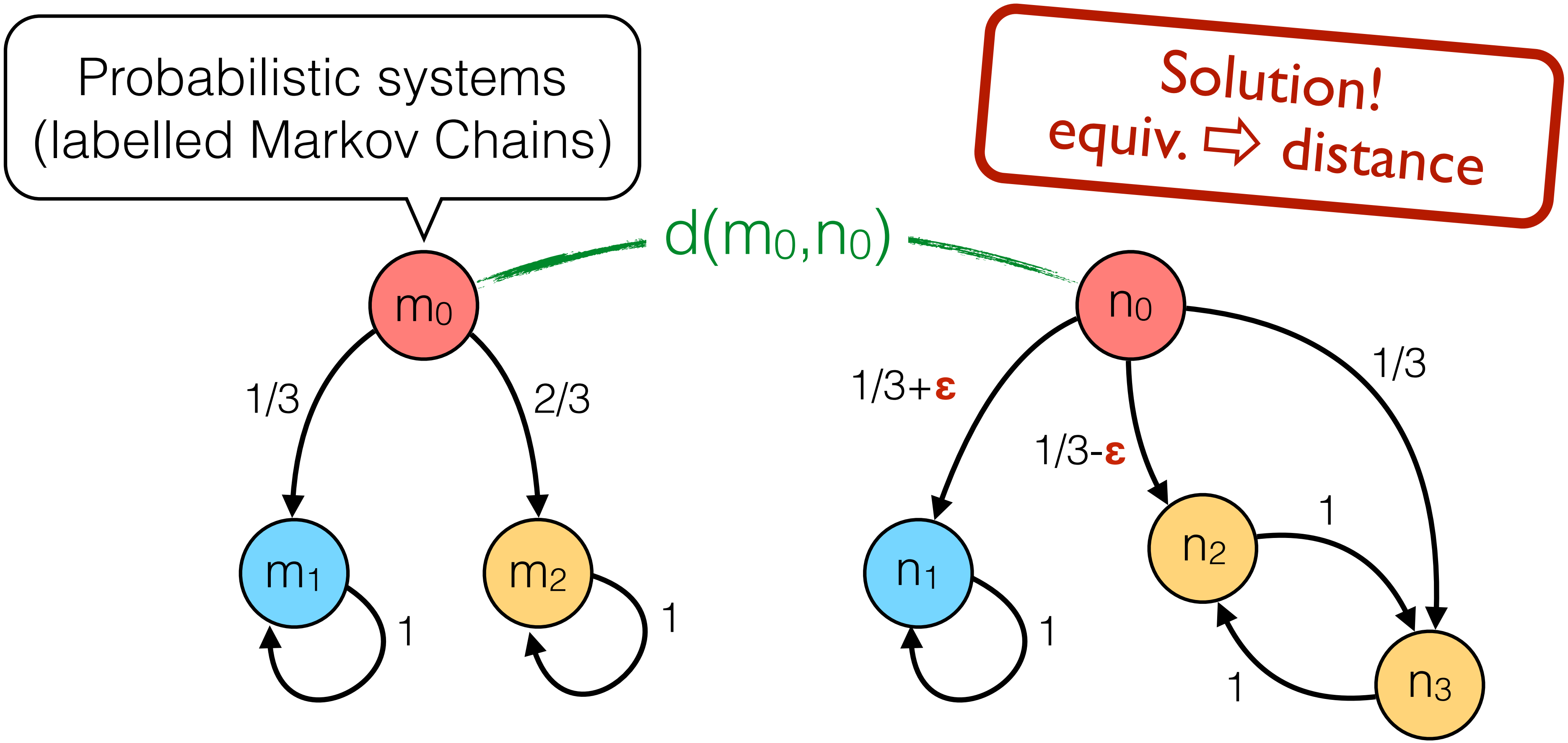
A fundamental problem

Jou-Smolka'90 observed that behavioral equivalences are not robust for systems with real-valued data



A fundamental problem

Jou-Smolka'90 observed that behavioral equivalences are not robust for systems with real-valued data



Metric-based Approximate Minimization

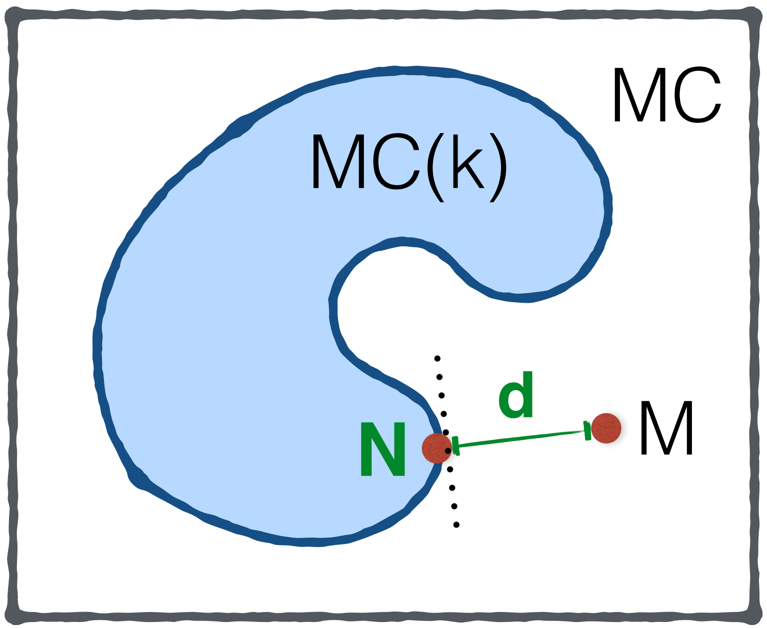
Closest Bounded Approximant (CBA)

Minimum Significant Approximant Bound (MSAB)

Metric-based Approximate Minimization

Closest Bounded Approximant (CBA)

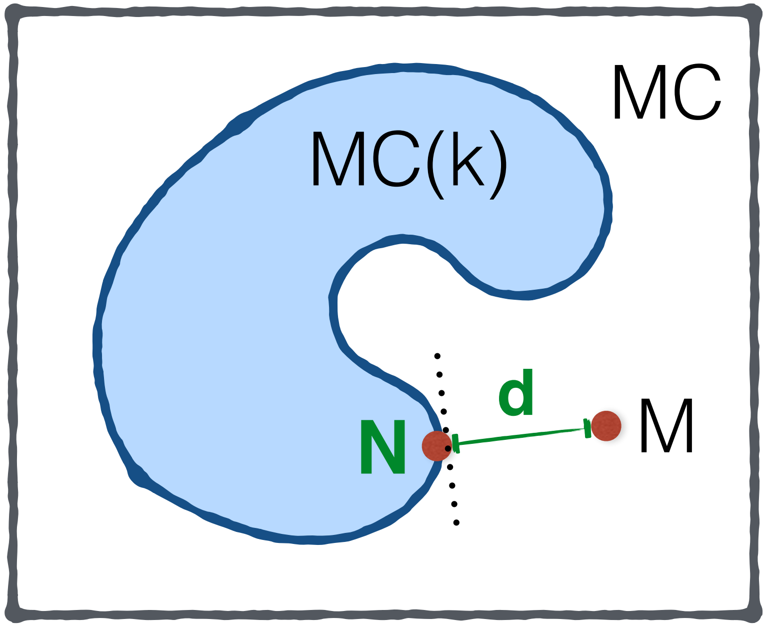
Minimum Significant Approximant Bound (MSAB)



Metric-based Approximate Minimization

Closest Bounded Approximant (CBA)

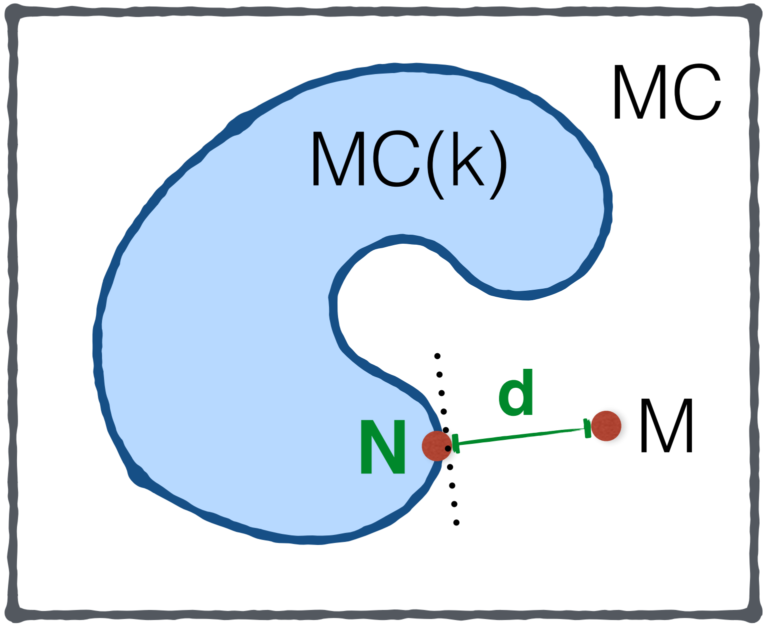
Minimum Significant Approximant Bound (MSAB)



minimize d

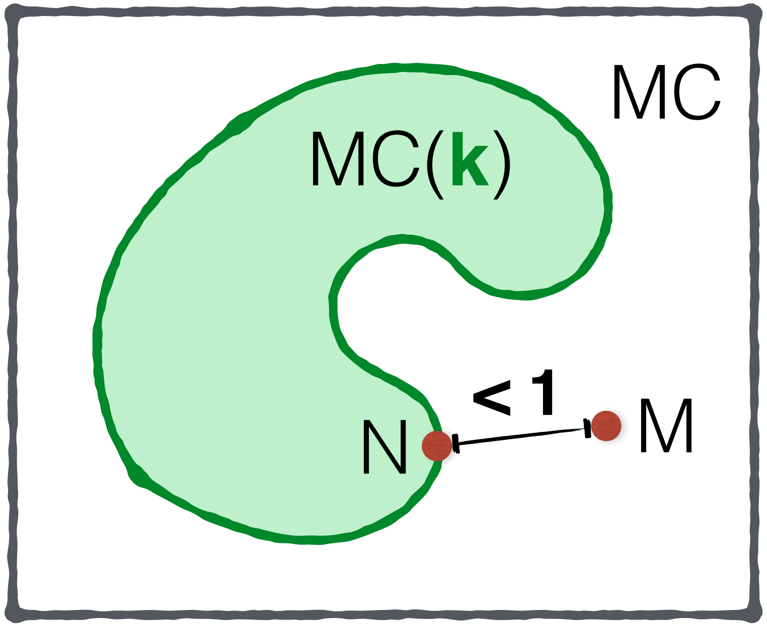
Metric-based Approximate Minimization

Closest Bounded Approximant (CBA)



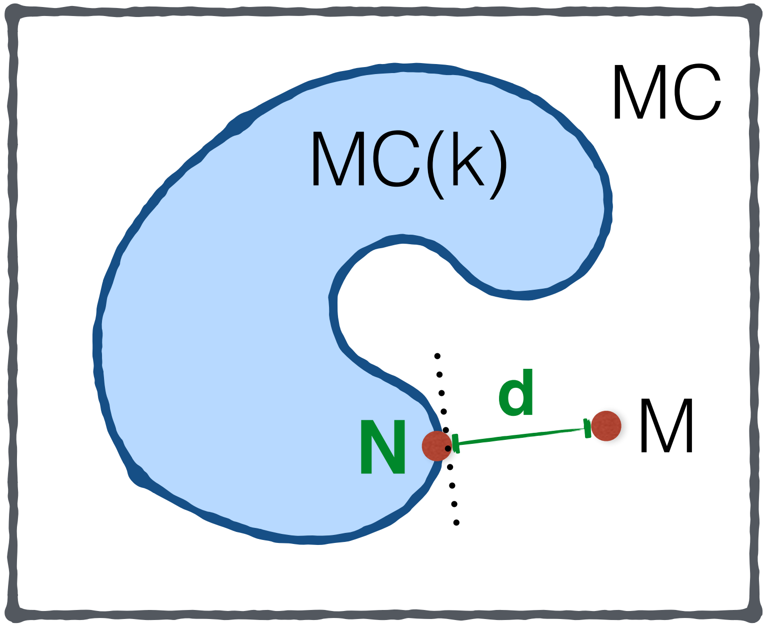
minimize d

Minimum Significant Approximant Bound (MSAB)



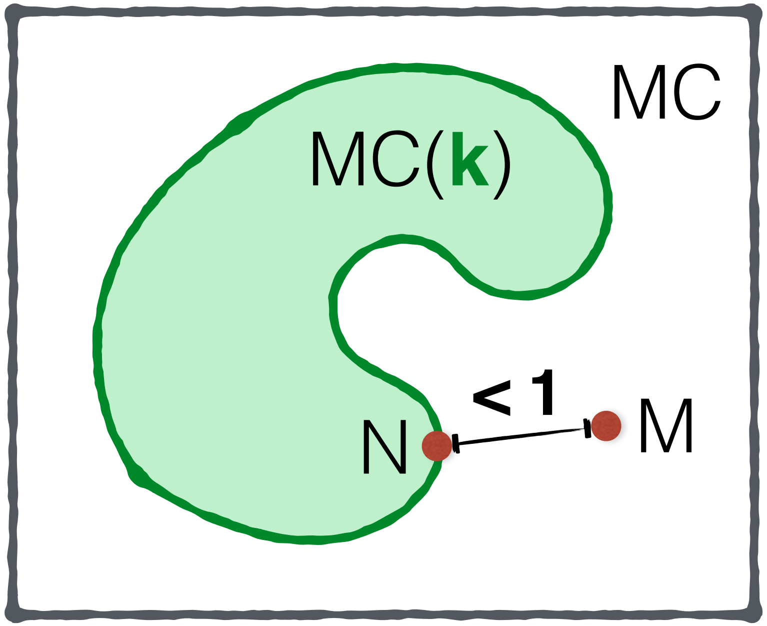
Metric-based Approximate Minimization

Closest Bounded Approximant (CBA)



minimize d

Minimum Significant Approximant Bound (MSAB)

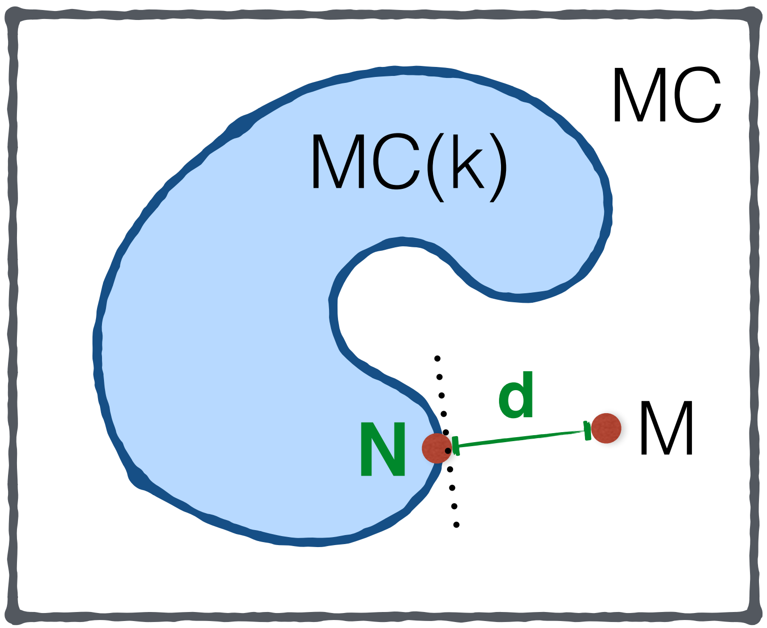


minimize k

"To study the complexity of an optimization problem one has to look at its decision variant"

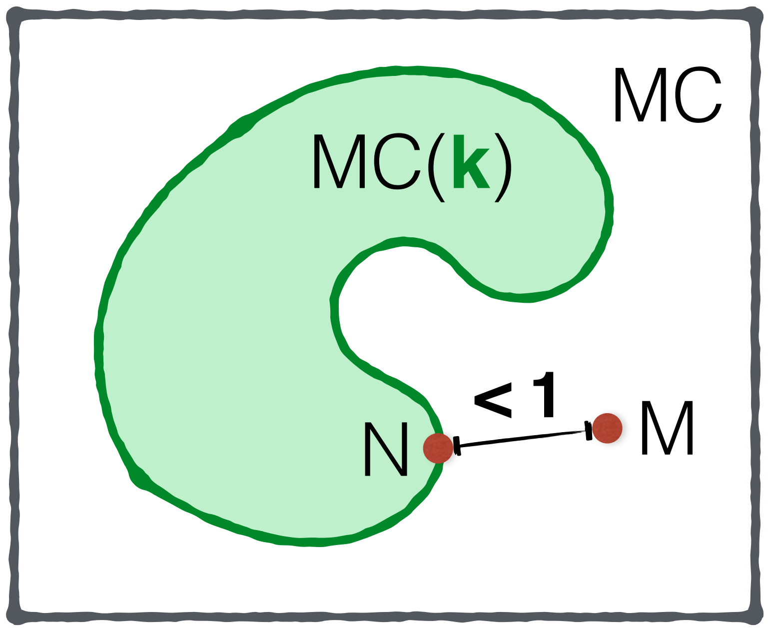
(C. Papadimitriou)

Closest Bounded Approximant (CBA)



minimize d

Minimum Significant Approximant Bound (MSAB)

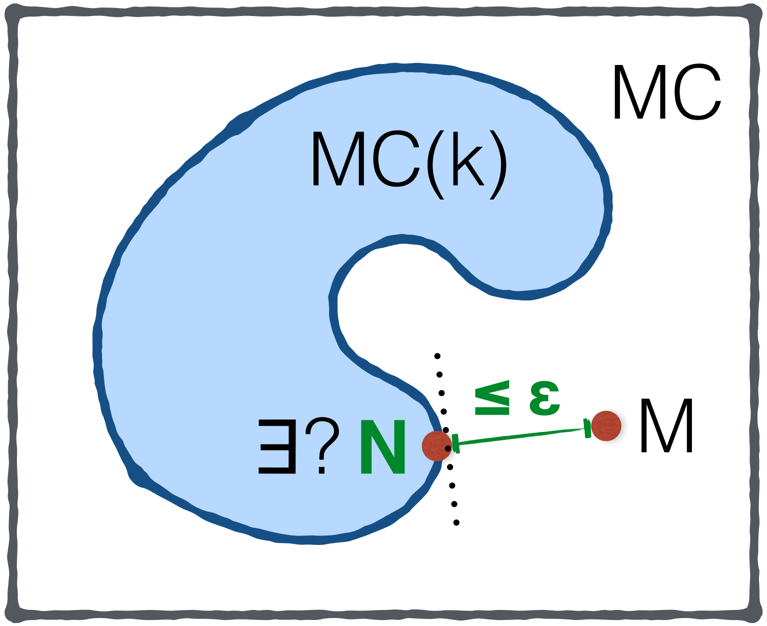


minimize k

"To study the complexity of an optimization problem one has to look at its decision variant"

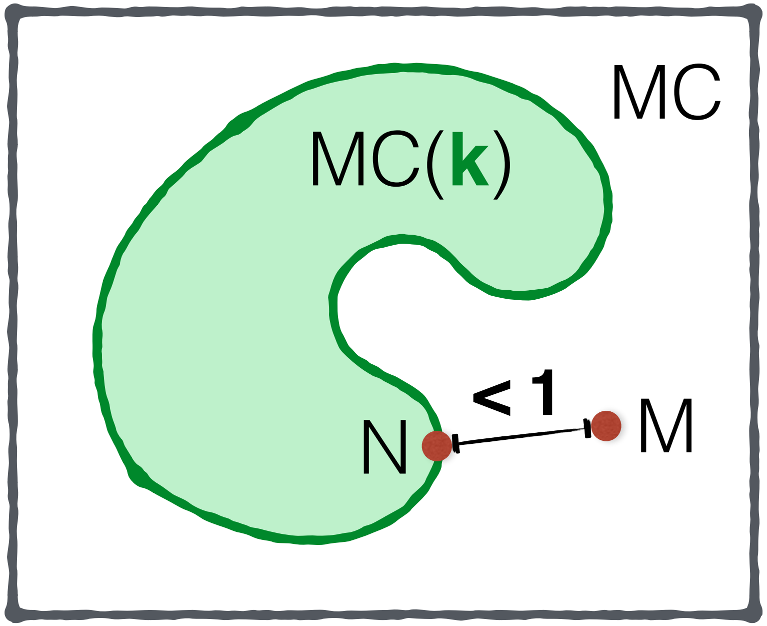
(C. Papadimitriou)

Bounded Approximant (BA)



given ε

Minimum Significant Approximant Bound (MSAB)

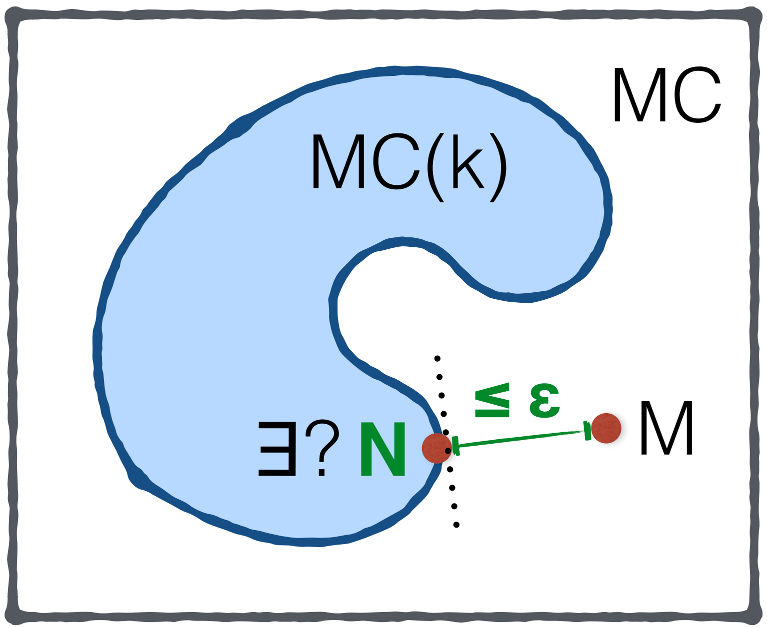


minimize k

"To study the complexity of an optimization problem one has to look at its decision variant"

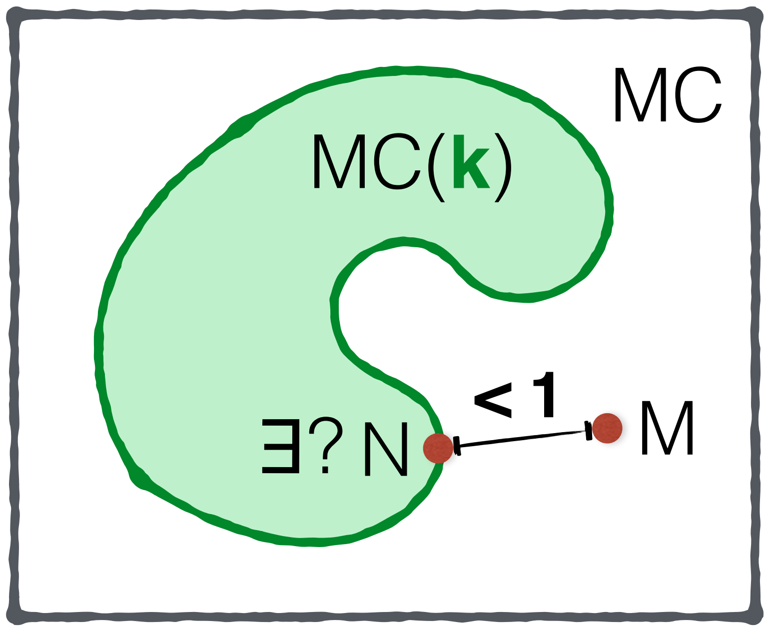
(C. Papadimitriou)

Bounded Approximant (BA)



given ε

Significant Bounded Approximant (SBA)



given k

What distance on MCs?

a.k.a. Kantorovich distance

(λ-discounted) **Probabilistic Bisimilarity distance** of Desharnais et al. —denoted d_λ

What distance on MCs?

a.k.a. Kantorovich distance

(λ-discounted) **Probabilistic Bisimilarity distance**
of Desharnais et al. —denoted d_λ

Theorem (Desharnais et. al 99)

$$m \sim n \quad \text{iff} \quad d_\lambda(m,n) = 0$$

What distance on MCs?

a.k.a. Kantorovich distance

(λ-discounted) **Probabilistic Bisimilarity distance**
of Desharnais et al. —denoted d_λ

Theorem (Desharnais et. al 99)

$$m \sim n \quad \text{iff} \quad d_\lambda(m,n) = 0$$

Theorem (Chen, van Breugel, Worrell 12)

The probabilistic bisimilarity distance
can be computed in polynomial time

Relation with Model Checking

Theorem (Chen, van Breugel, Worrell 12)

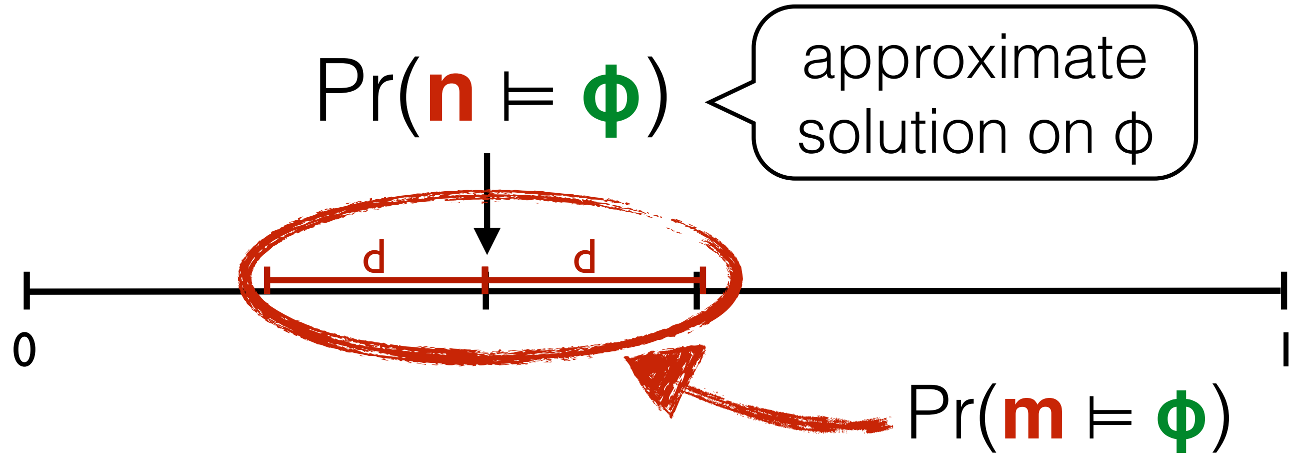
For all $\phi \in LTL \quad | Pr(m \models \phi) - Pr(n \models \phi) | \leq d_1(m,n)$

Relation with Model Checking

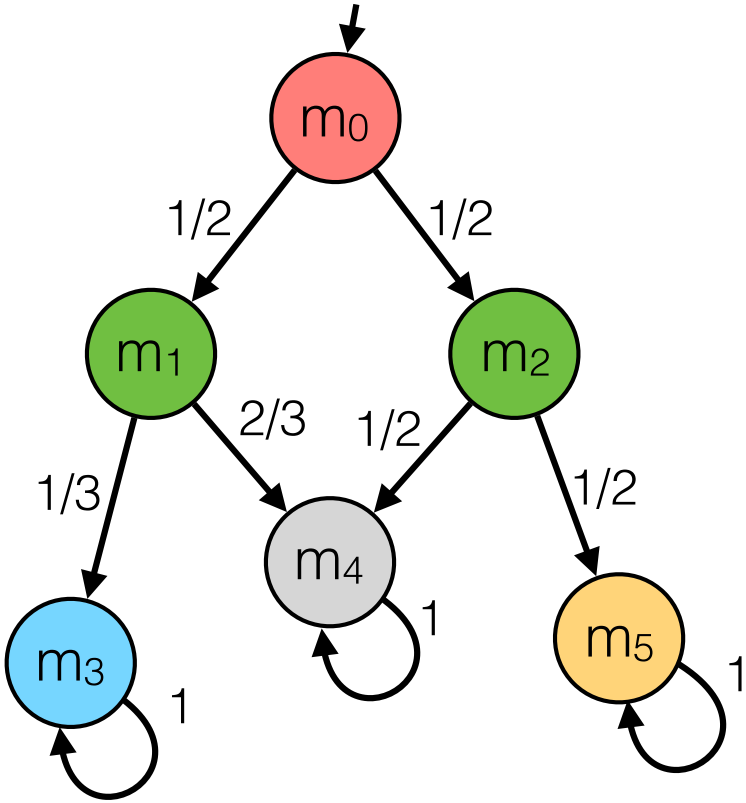
Theorem (Chen, van Breugel, Worrell 12)

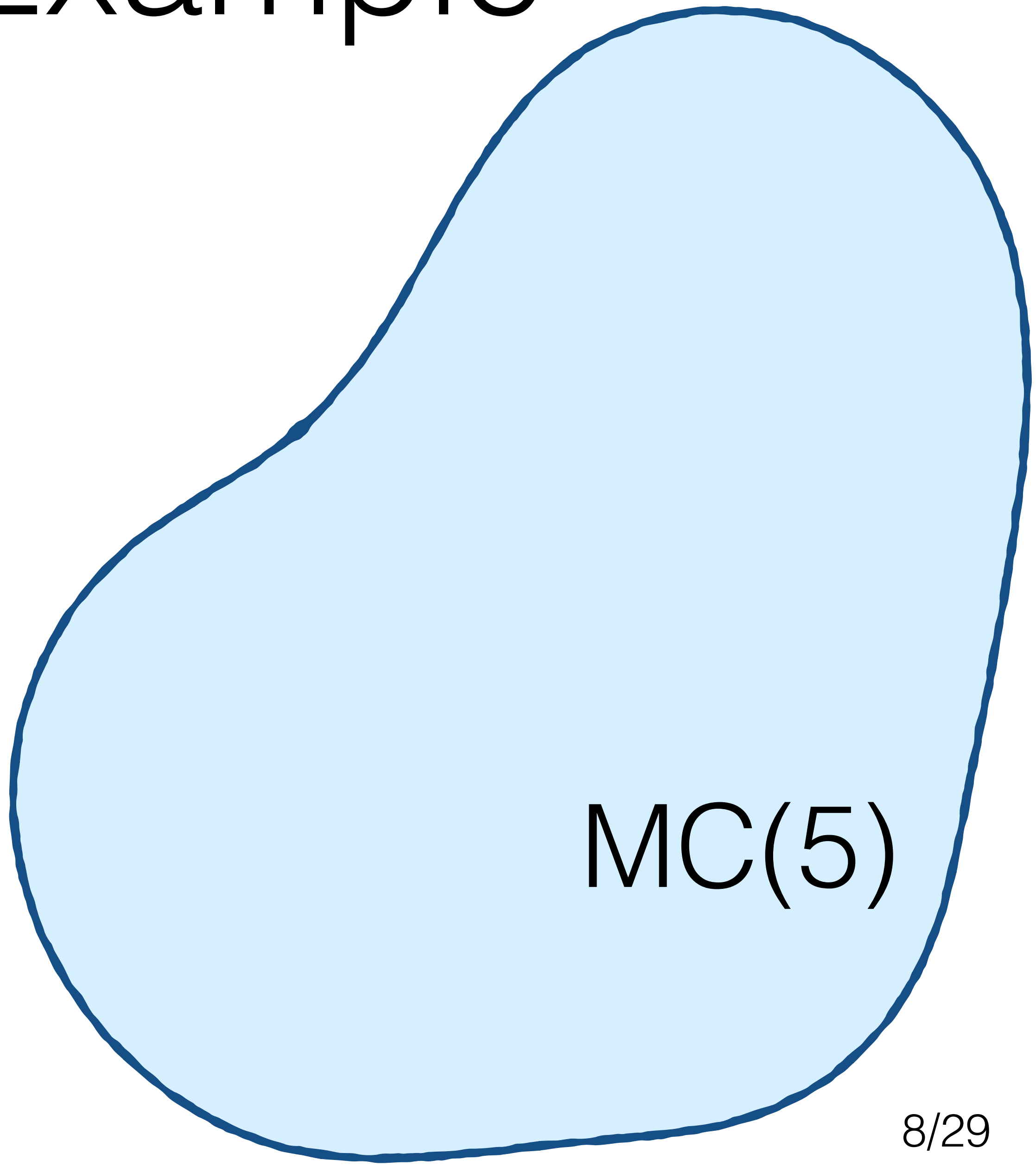
For all $\phi \in$ LTL $\quad | \Pr(m \vDash \phi) - \Pr(n \vDash \phi) | \leq d_1(m,n)$

…imagine that $|M| \gg |N|$, we can use N in place of M



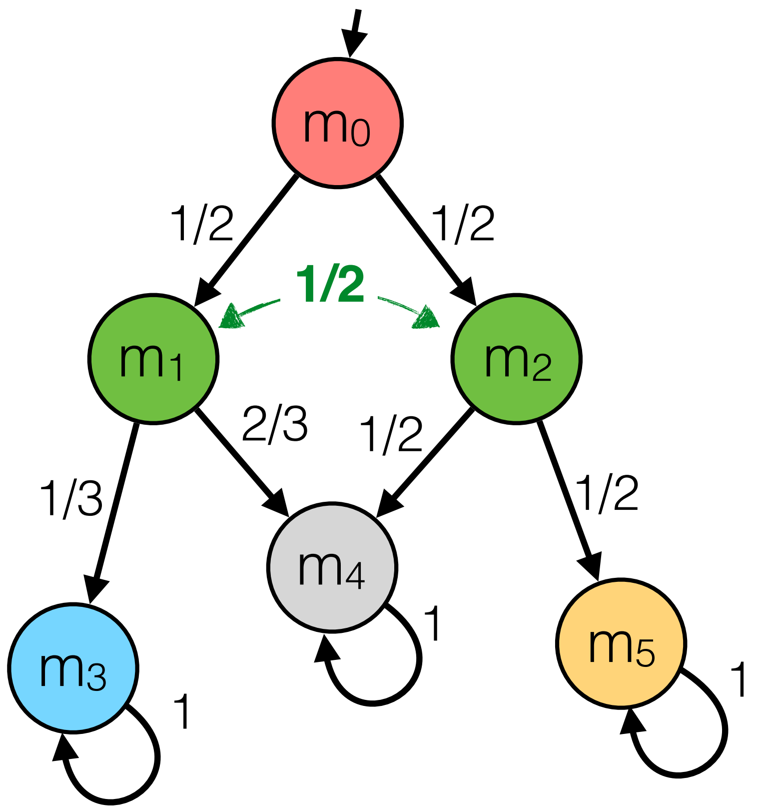
CBA: Example*

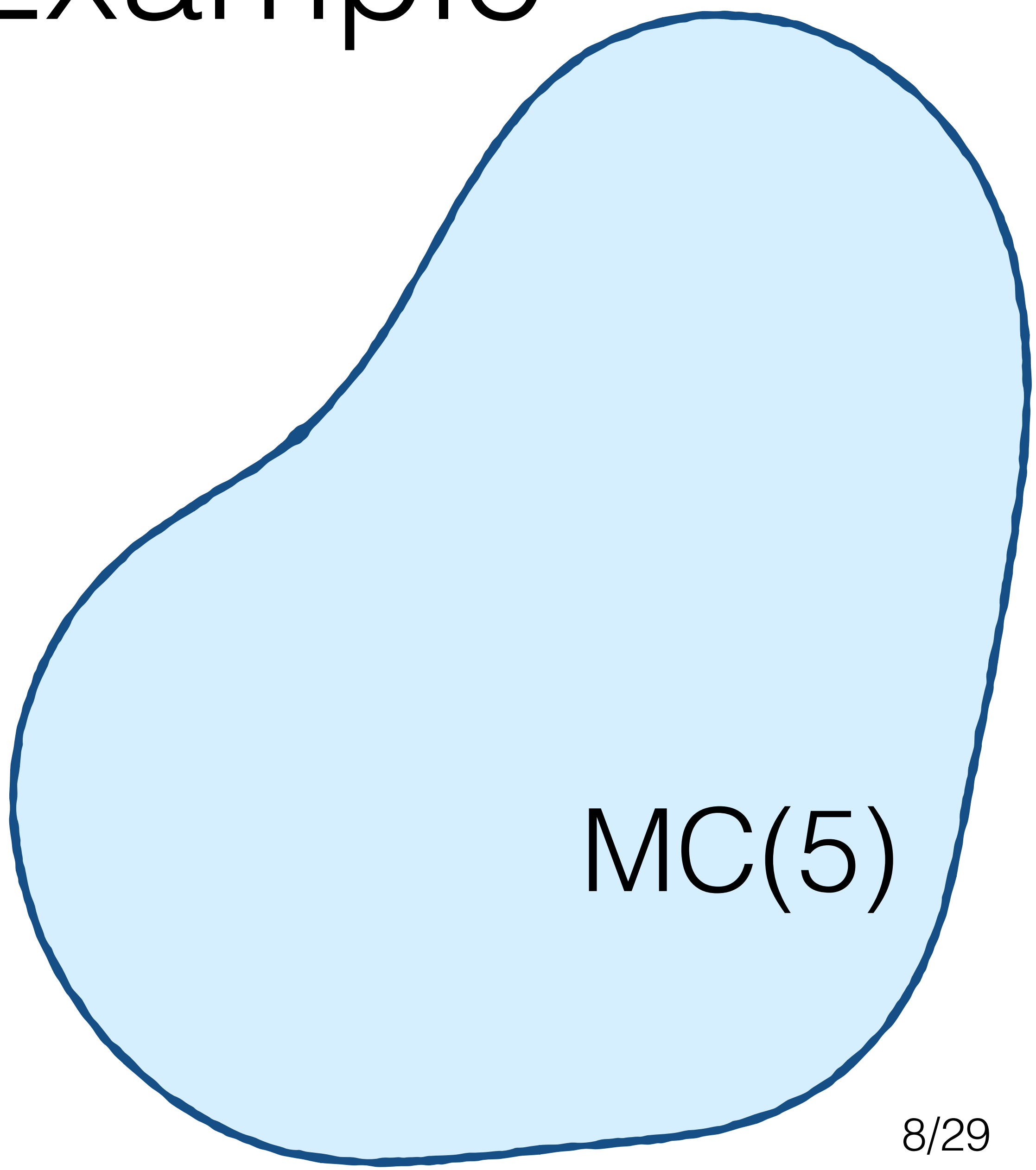




(*) With respect to the undiscounted probabilistic bisimilarity distance d_1

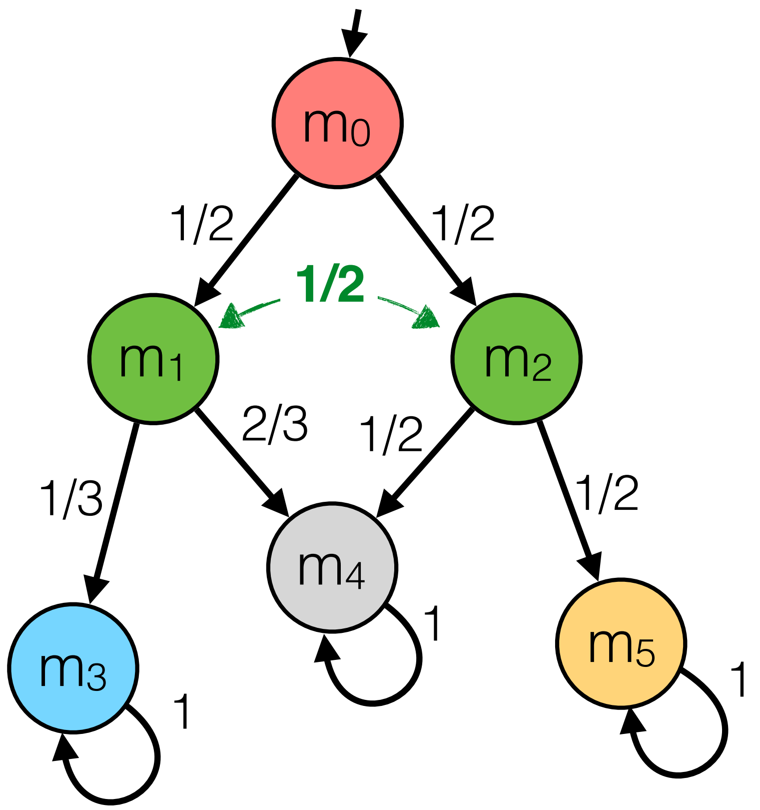
CBA: Example*

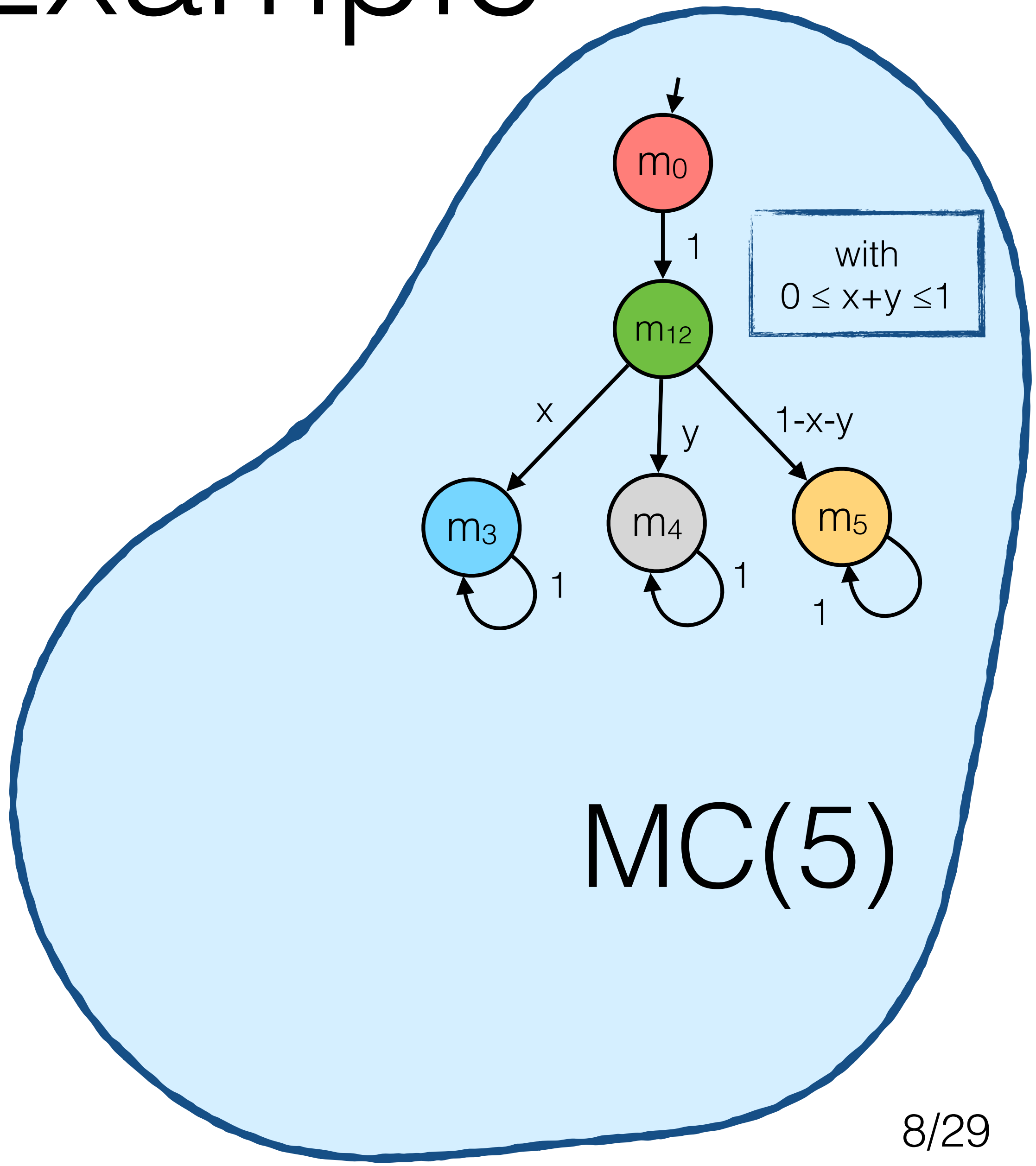




(*) With respect to the undiscounted probabilistic bisimilarity distance d_1

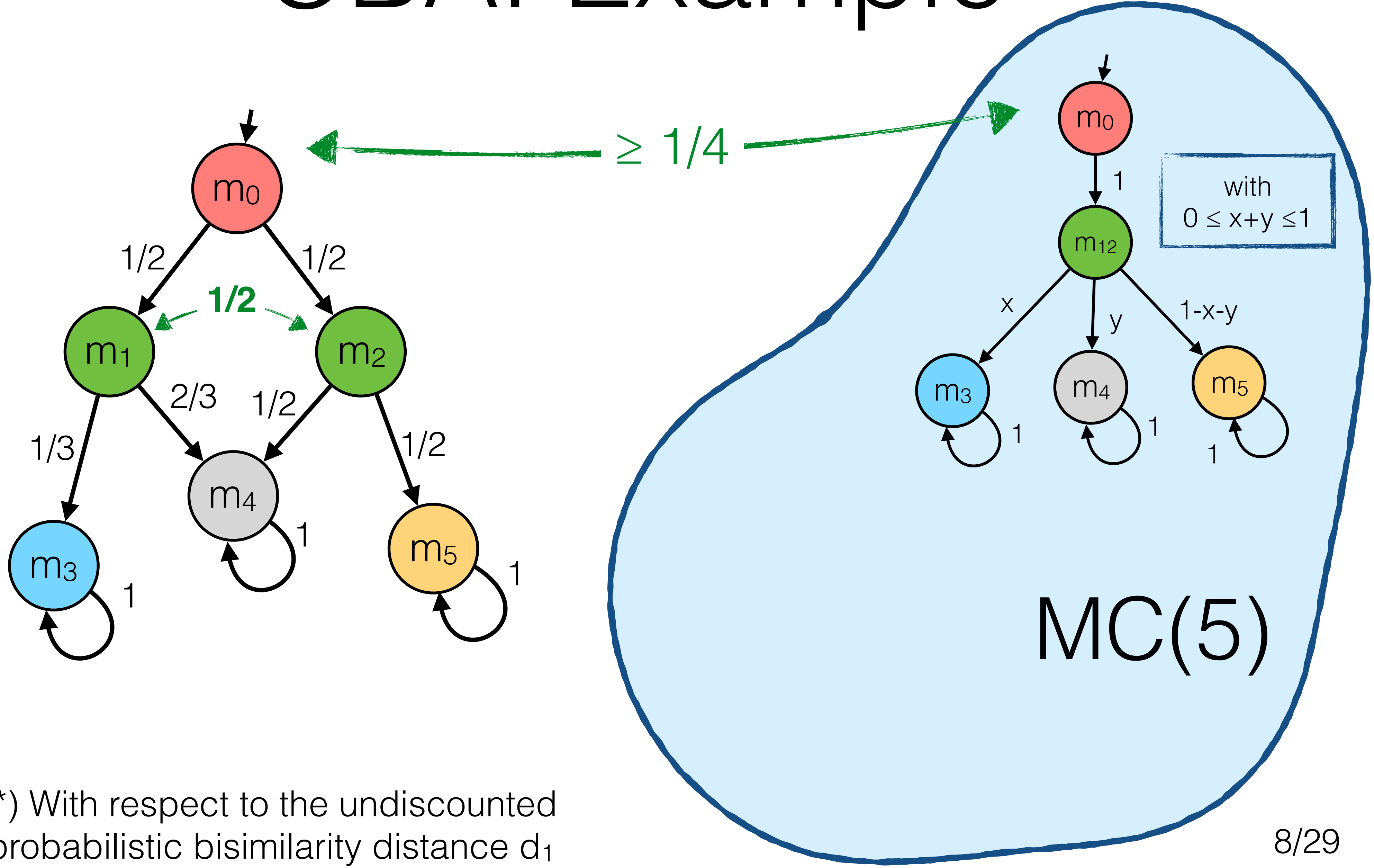
CBA: Example*





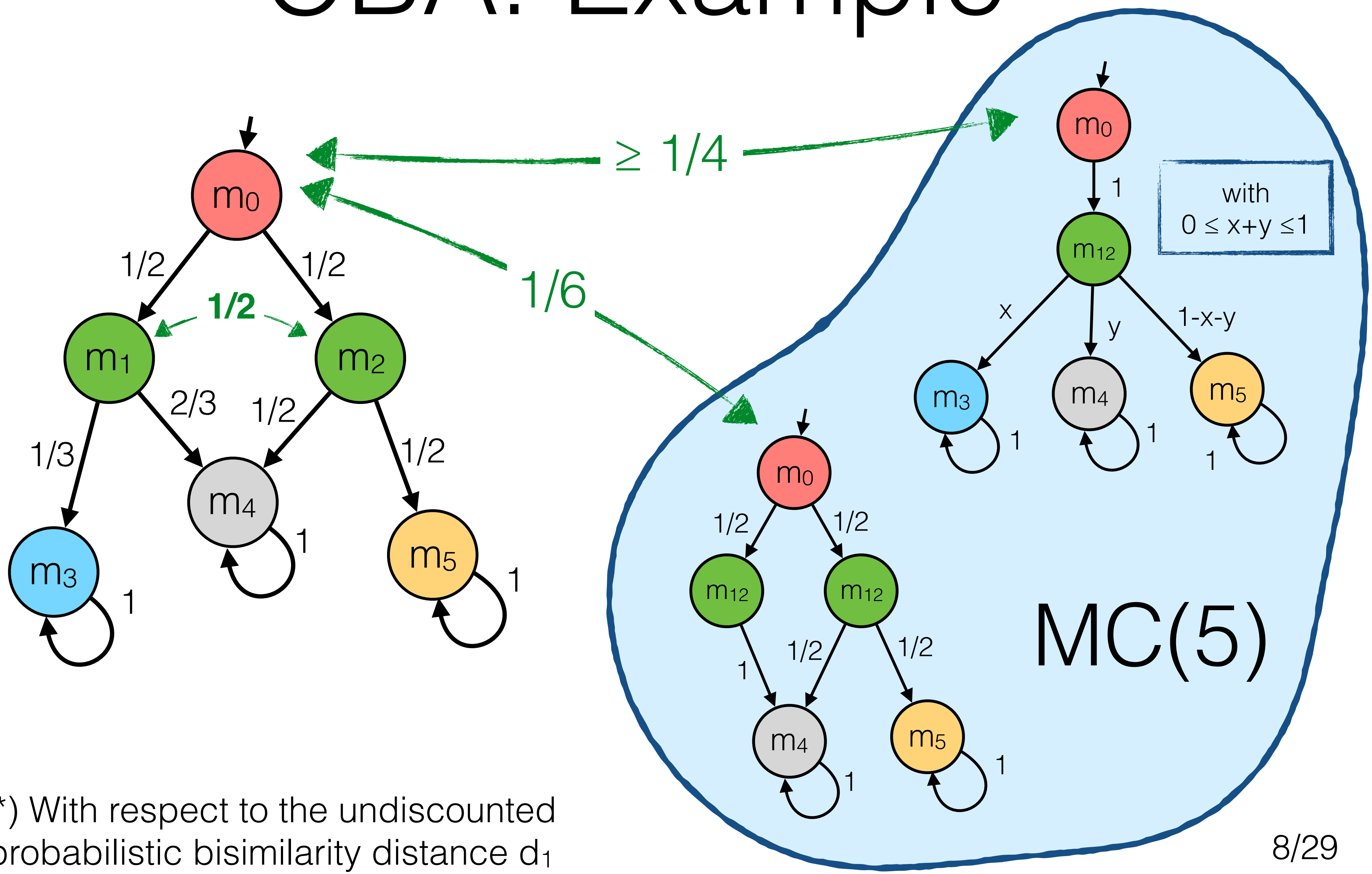
(*) With respect to the undiscounted probabilistic bisimilarity distance d_1

CBA: Example*



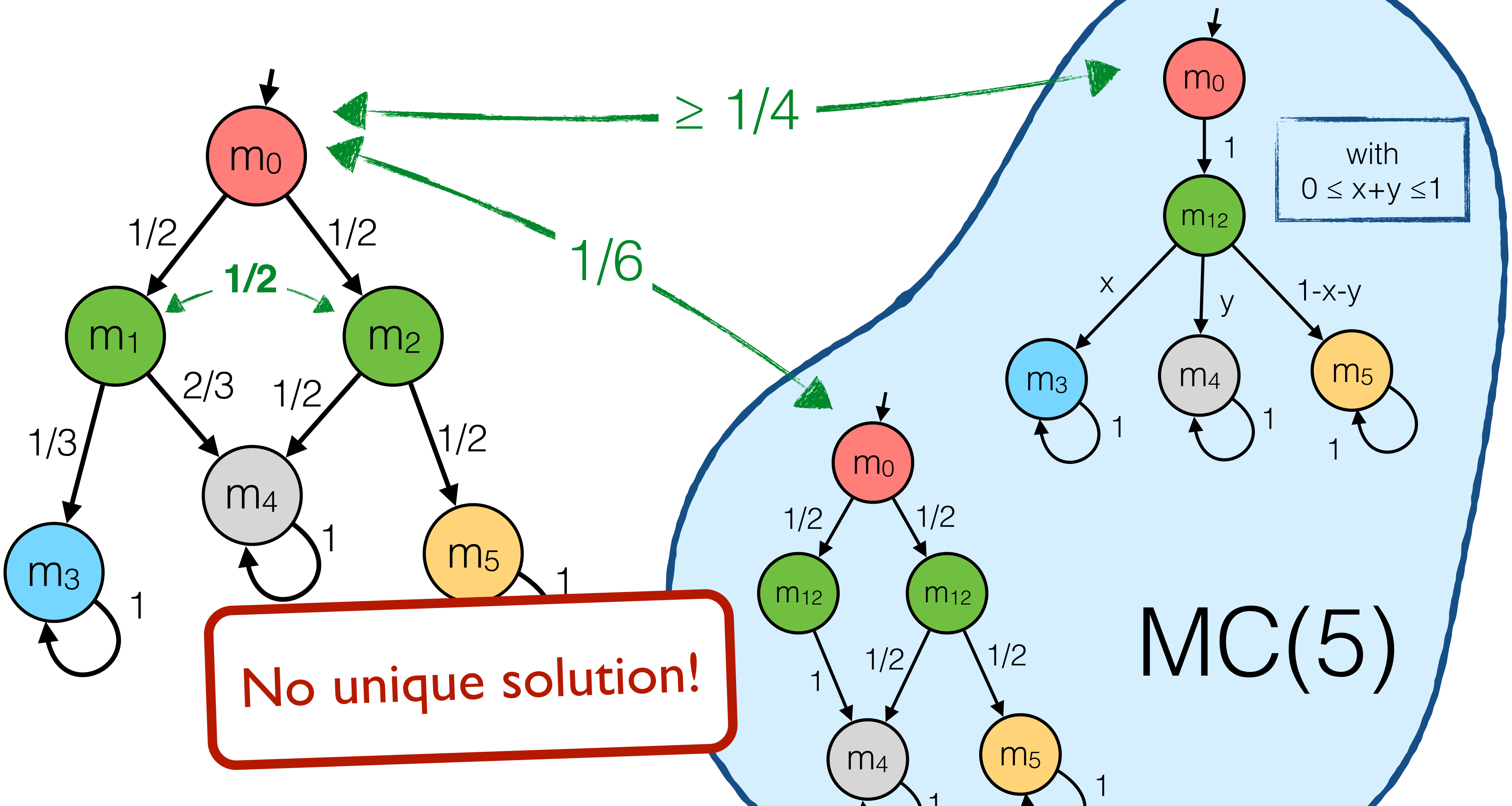
(*) With respect to the undiscounted probabilistic bisimilarity distance d_1

CBA: Example*



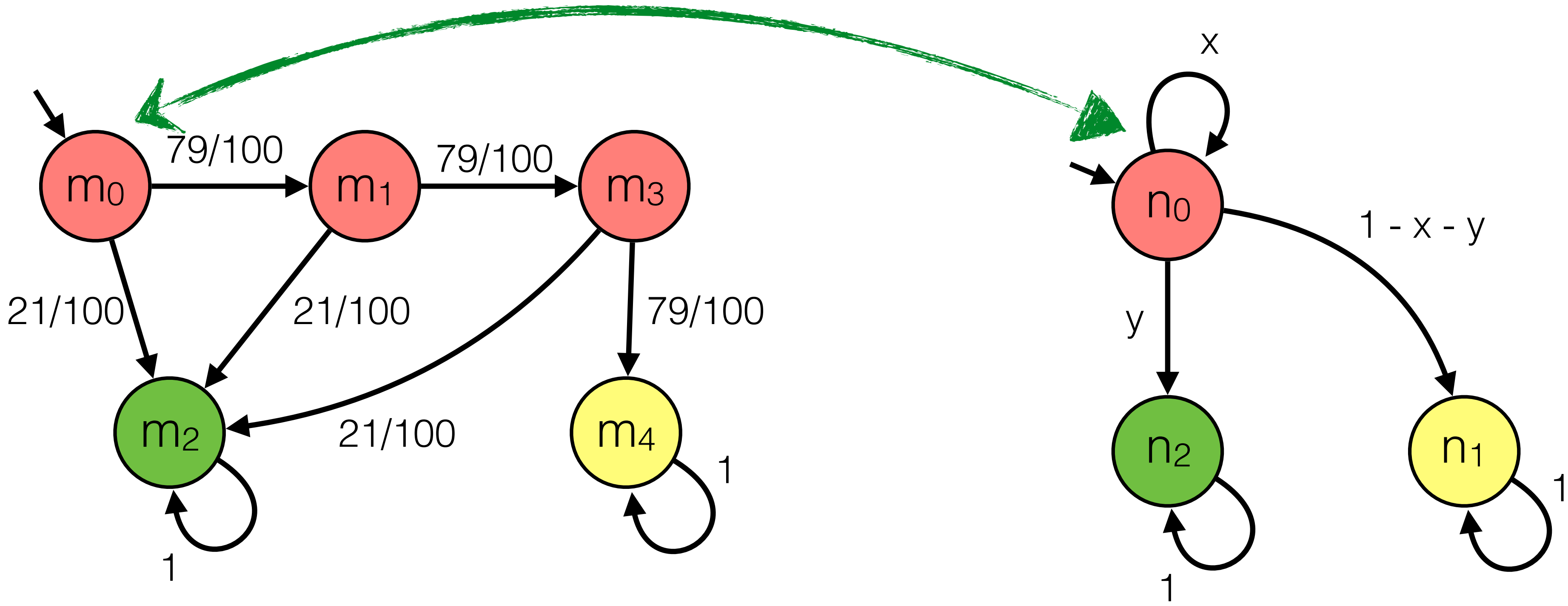
(*) With respect to the undiscounted probabilistic bisimilarity distance d_1

CBA: Example*



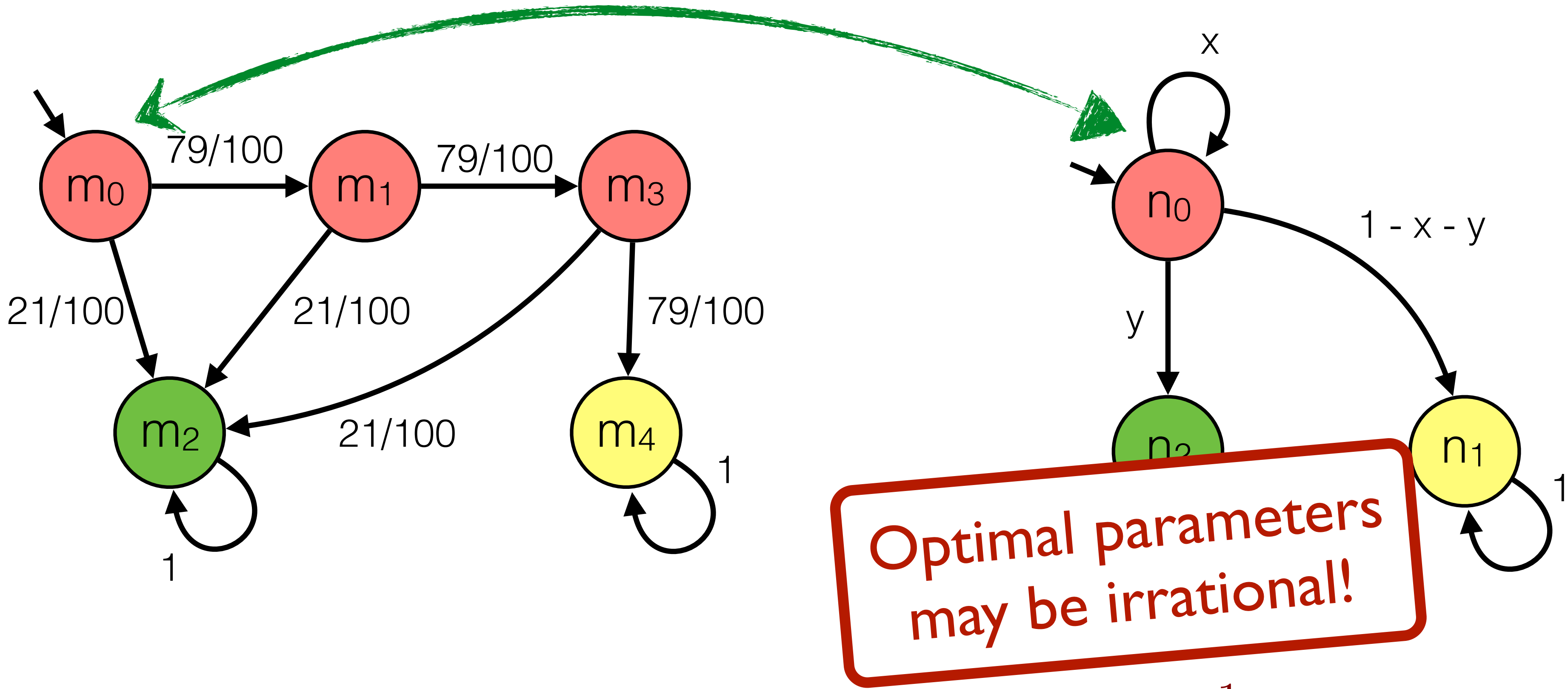
(*) With respect to the undiscounted probabilistic bisimilarity distance d_1

CBA: Example*



(*) With respect to the undiscounted probabilistic bisimilarity distance d_1

CBA: Example*

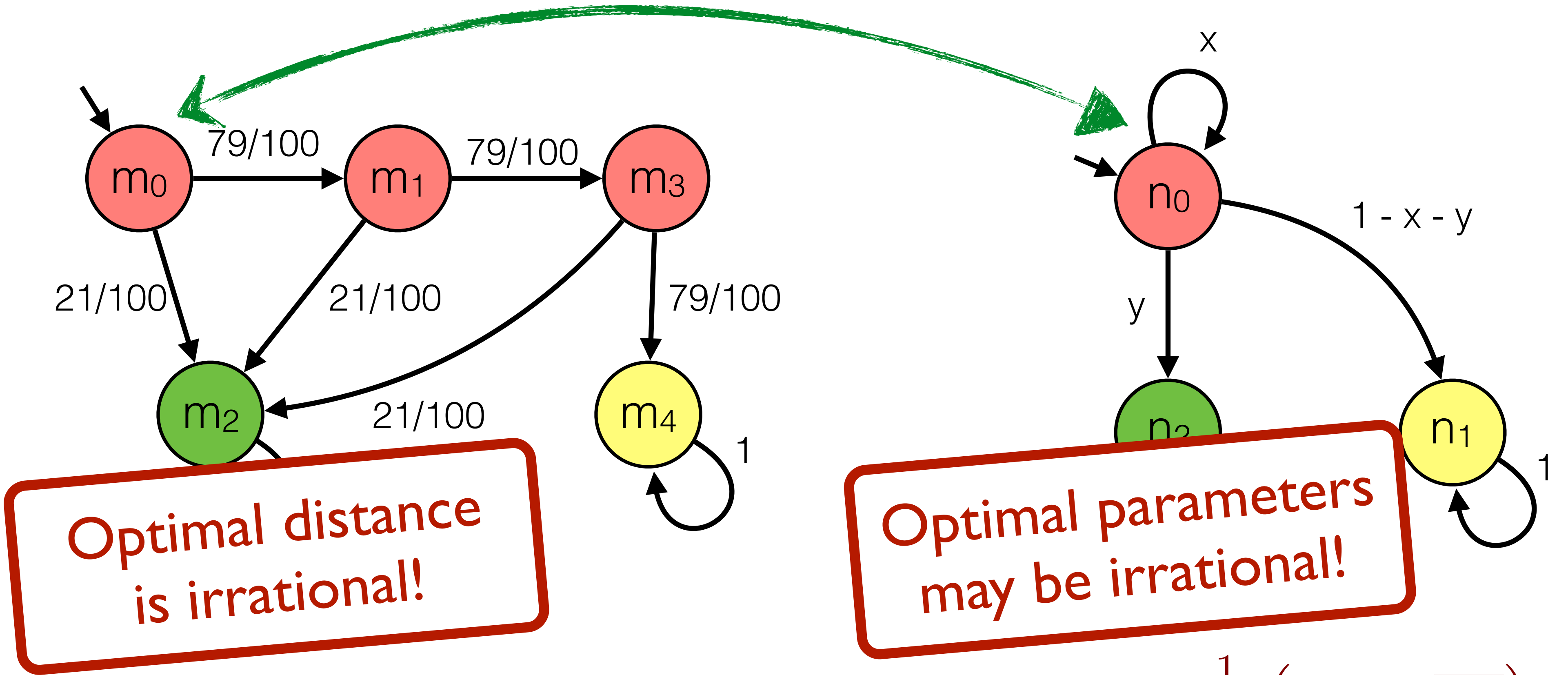


$$x = \frac{1}{30}\left(10 + \sqrt{163}\right)$$

$$y = \frac{21}{200}$$

(*) With respect to the undiscounted probabilistic bisimilarity distance d_1

CBA: Example*



$$\delta(m_0, n_0) = \frac{436}{675} - \frac{163\sqrt{163}}{13500} \approx 0.49$$

$$x = \frac{1}{30}\left(10 + \sqrt{163}\right)$$

$$y = \frac{21}{200}$$

(*) With respect to the undiscounted probabilistic bisimilarity distance d_1

Our Contributions

Our Contributions

Characterizations + COMPLEXITY results:

Our Contributions

Characterizations + COMPLEXITY results:

1. Closest Bounded Approximant (CBA)
   encoded as a bilinear program

Our Contributions

Characterizations + COMPLEXITY results:

1. Closest Bounded Approximant (CBA)
   encoded as a bilinear program
2. Bounded Approximant (BA)
   PSPACE & NP-hard for all $\lambda \in (0,1]$

Our Contributions

Characterizations + COMPLEXITY results:

1. Closest Bounded Approximant (CBA)
   encoded as a bilinear program
2. Bounded Approximant (BA)
   PSPACE & NP-hard for all $\lambda \in (0,1]$
3. Significant Bounded Approximant (SBA)
   NP-complete for $\lambda = 1$

Our Contributions

Theoretical

Characterizations + COMPLEXITY results:

1. Closest Bounded Approximant (CBA)
   encoded as a bilinear program
2. Bounded Approximant (BA)
   PSPACE & NP-hard for all $\lambda \in (0,1]$
3. Significant Bounded Approximant (SBA)
   NP-complete for $\lambda = 1$

Our Contributions

Theoretical

Characterizations + COMPLEXITY results:

1. Closest Bounded Approximant (CBA)
   encoded as a bilinear program
2. Bounded Approximant (BA)
   PSPACE & NP-hard for all $\lambda \in (0,1]$
3. Significant Bounded Approximant (SBA)
   NP-complete for $\lambda = 1$

Practical

We proposed an EM-like method to obtain a sub-optimal approximants

Talk Outline

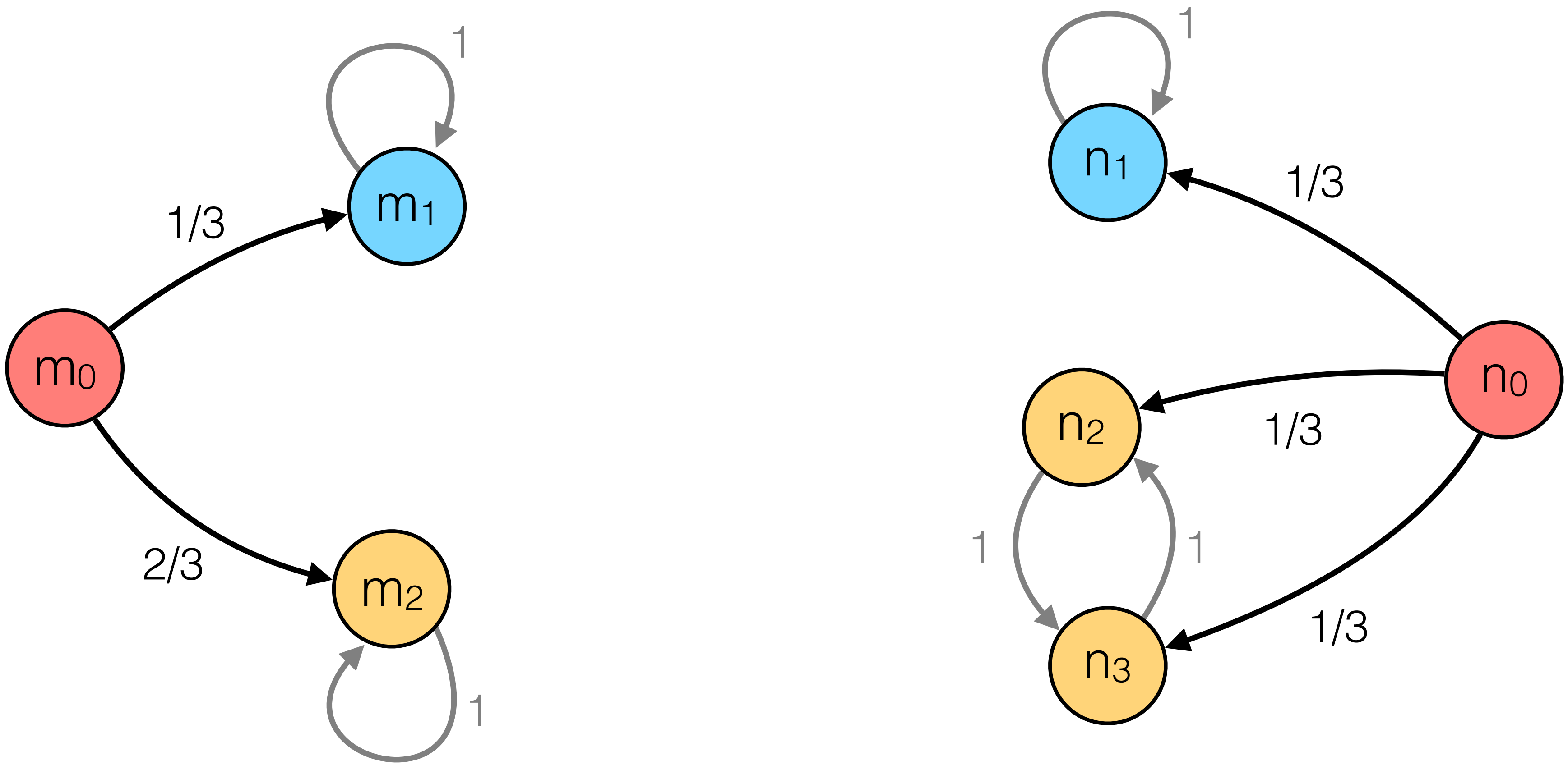
★ **Probabilistic bisimilarity distance**

- fixed point characterization (Kantorovich oper.)

★ **Metric-based Optimal Approximate Minimization**

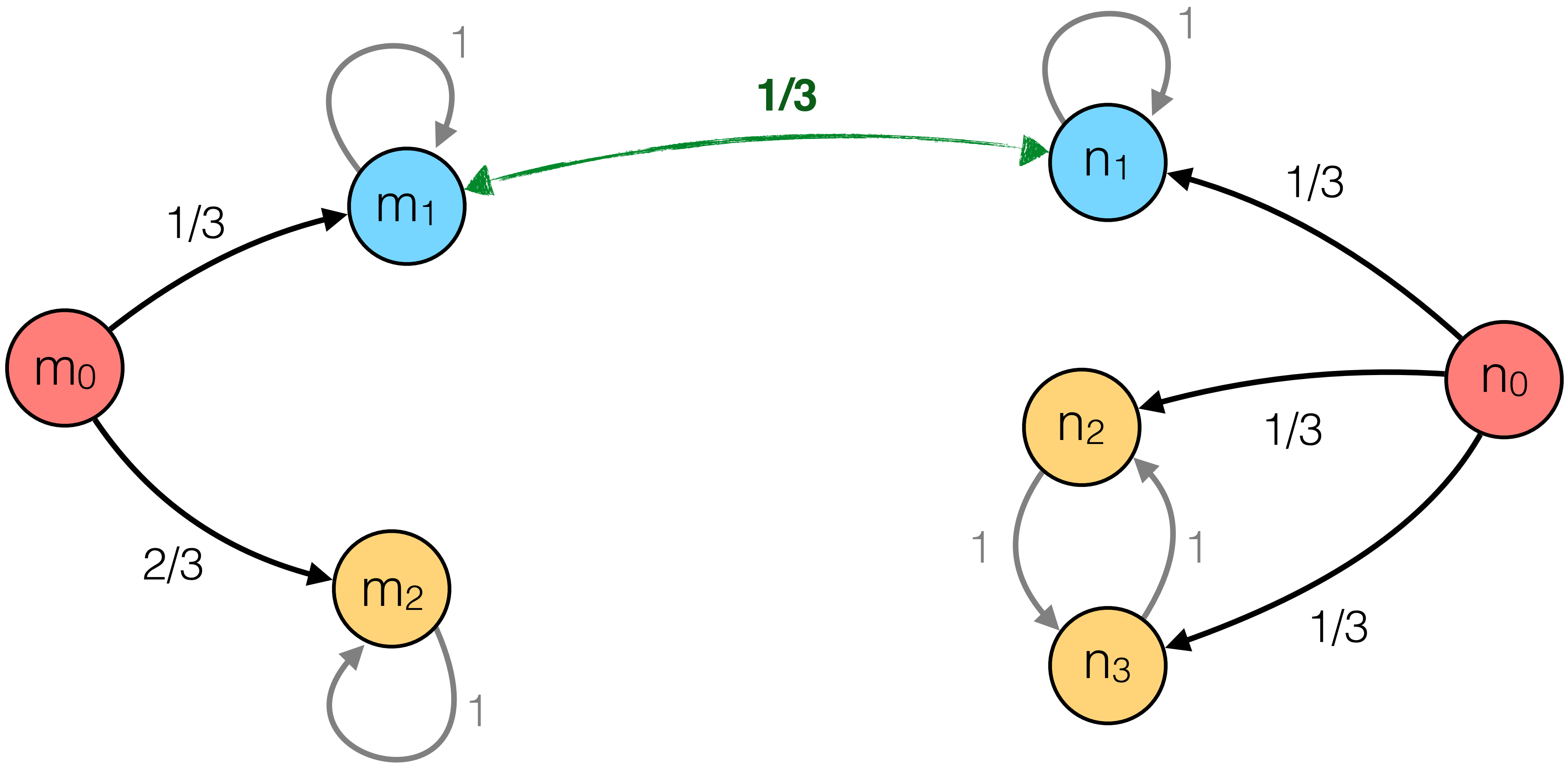
- *Closest Bounded Approximant* (CBA)
  — bilinear characterization (+ complexity)
- *Minimum Significant Approximant Bound* (MSAB)
  — characterization (+ complexity)
- *Expectation Maximization-like algorithm*
  — 2 heuristics + experimental results

Probabilistic bisimulation



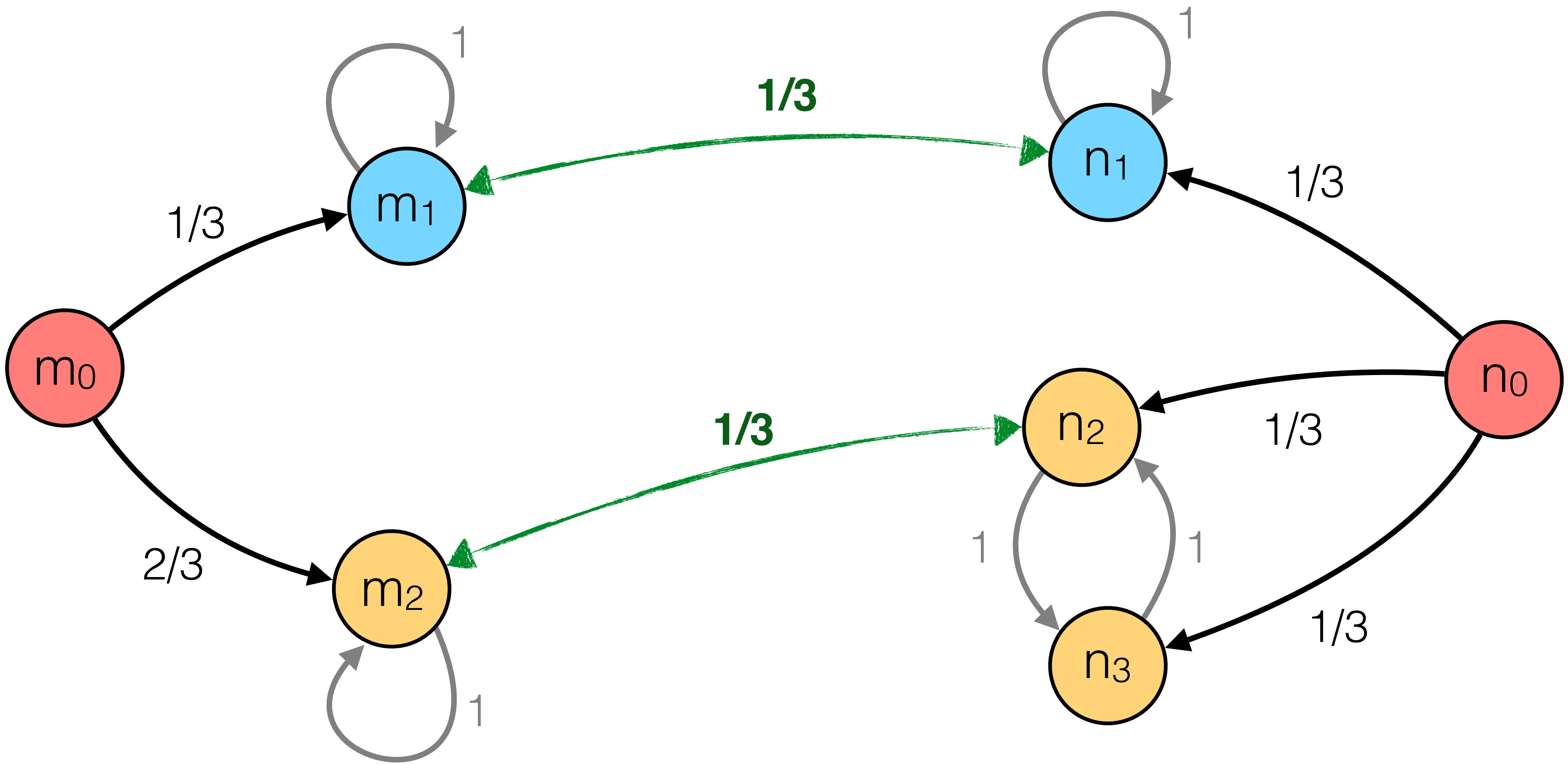
It tries to match the behaviors “quantitatively”

Probabilistic bisimulation



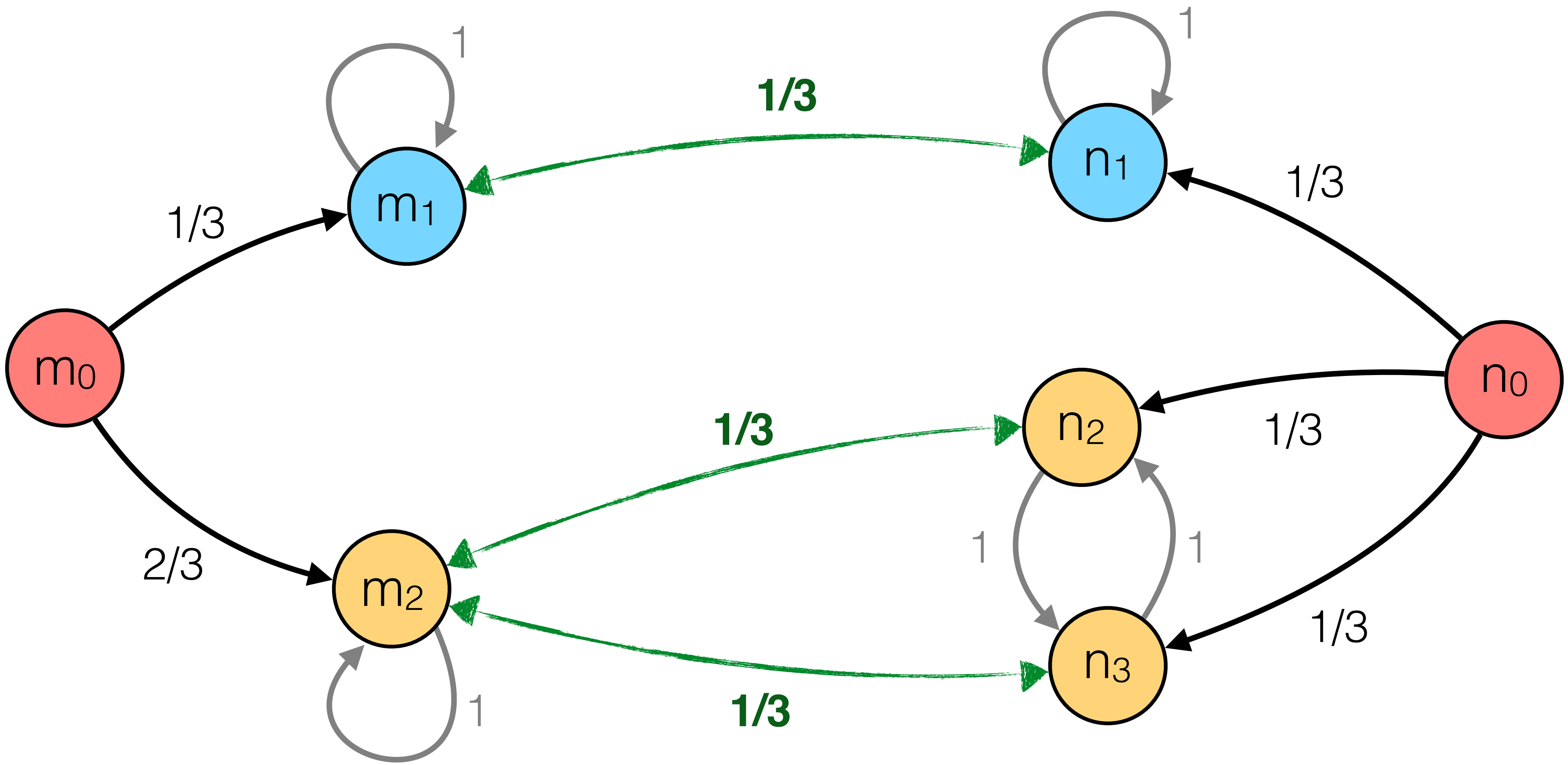
It tries to match the behaviors “quantitatively”

Probabilistic bisimulation



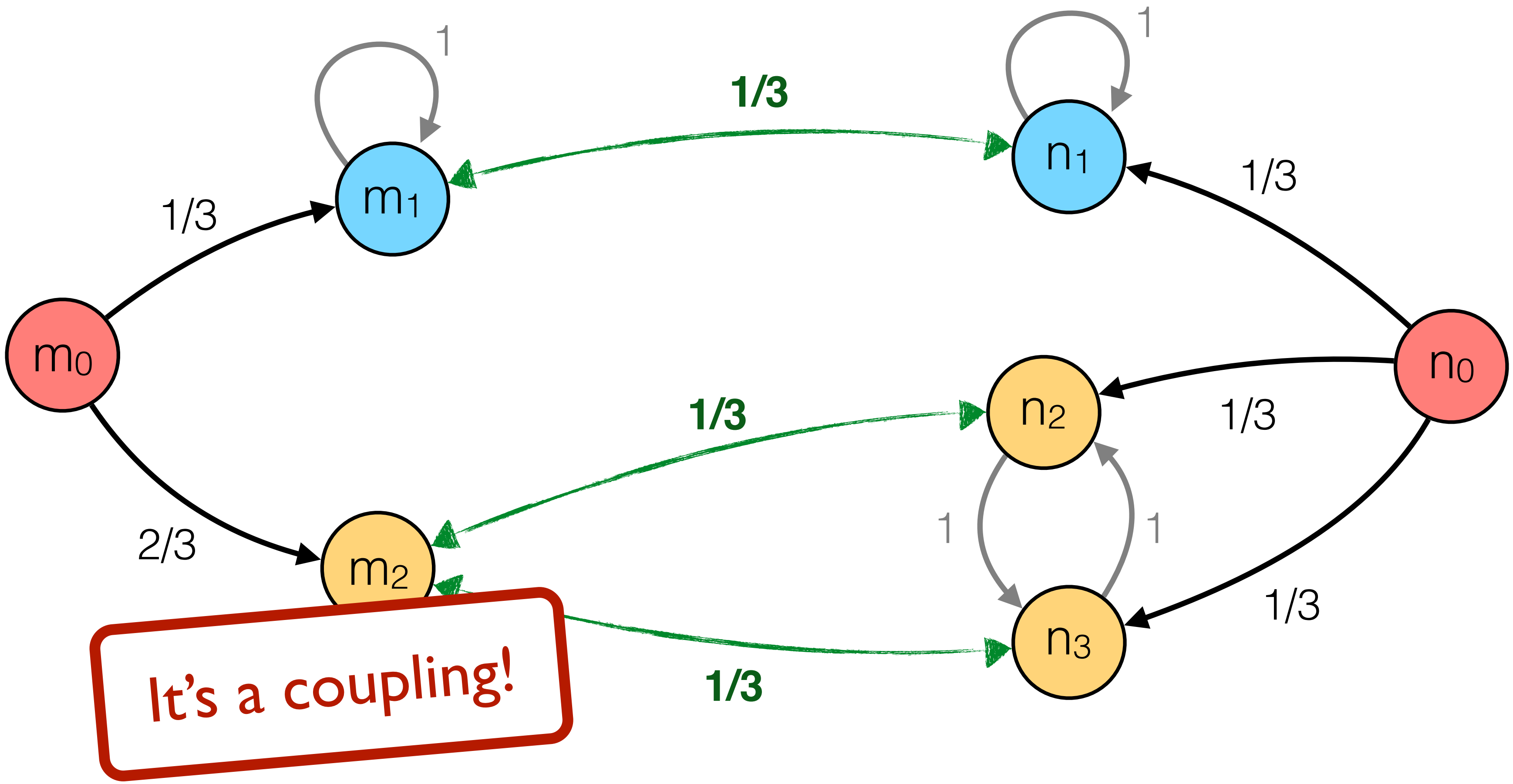
It tries to match the behaviors "quantitatively"

Probabilistic bisimulation



It tries to match the behaviors "quantitatively"

Probabilistic bisimulation



It tries to match the behaviors "quantitatively"

Coupling

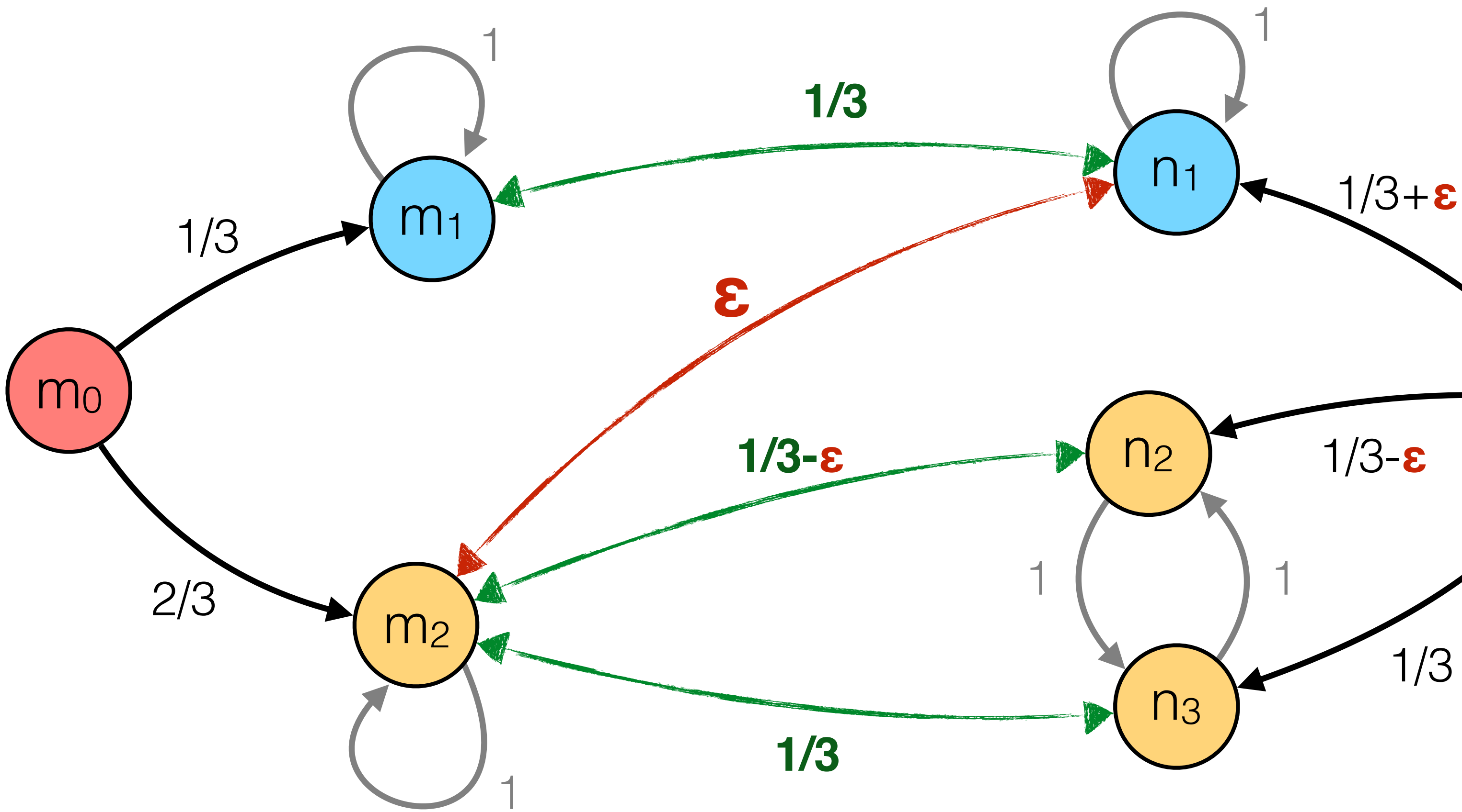
Definition (W. Doeblin 36)

A *coupling* of a pair (μ,ν) of probability distributions on M is a distribution ω on M×M such that

- $\sum_{n \in M} \omega(m,n) = \mu(m)$ *(left marginal)*
- $\sum_{m \in M} \omega(m,n) = \nu(n)$ *(right marginal)*.

One can think of a coupling as a measure-theoretic relation between probability distribution

A quantitative generalization



$$\text{minimize} \sum_{u,v \in M} \omega(u,v)\, d(u,v)$$

A quantitative generalization of probabilistic bisimilarity

(Desharnais et al.'99 & Worrell-van Breugel'00)

The **λ**-discounted **probabilistic bisimilarity pseudometric** is the smallest $d_\lambda : M \times M \to [0,1]$ such that

$$d_\lambda(m,n) = \Gamma_\lambda(d_\lambda) = \begin{cases} 1 & \text{if } \ell(m) \neq \ell(n) \\ \lambda\, K(d_\lambda)(\tau(m),\tau(n)) & \text{otherwise} \end{cases}$$

A quantitative generalization of probabilistic bisimilarity

(Desharnais et al.'99 & Worrell-van Breugel'00)

The **λ**-discounted **probabilistic bisimilarity pseudometric** is the smallest $d_\lambda : M \times M \to [0,1]$ such that

labels

$$d_\lambda(m,n) = \Gamma_\lambda(d_\lambda) = \begin{cases} 1 & \text{if } \ell(m) \neq \ell(n) \\ \lambda\, K(d_\lambda)(\tau(m),\tau(n)) & \text{otherwise} \end{cases}$$

A quantitative generalization of probabilistic bisimilarity

(Desharnais et al.'99 & Worrell-van Breugel'00)

The **λ**-discounted **probabilistic bisimilarity pseudometric** is the smallest $d_\lambda : M \times M \to [0,1]$ such that

$$d_\lambda(m,n) = \Gamma_\lambda(d_\lambda) = \begin{cases} 1 & \text{if } \ell(m) \neq \ell(n) \\ \lambda\, K(d_\lambda)(\tau(m),\tau(n)) & \text{otherwise} \end{cases}$$

labels

transition probabilities

Kantorovich distance

$$K(d)(\mu,\nu) = \min_{\omega \in \Omega(\mu,\nu)} \sum_{u,v \in M} \omega(u,v)\, d(u,v)$$

Talk Outline

- ★ **Probabilistic bisimilarity distance**
  - fixed point characterization (Kantorovich oper.)

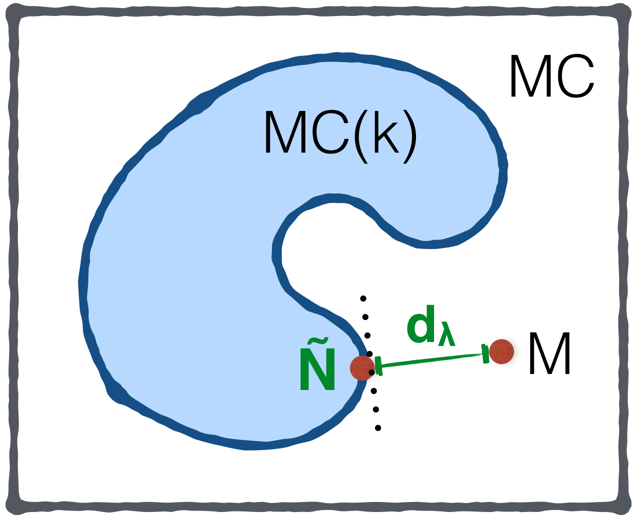
- ★ **Metric-based Optimal Approximate Minimization**
  - *Closest Bounded Approximant* (CBA)
    — bilinear characterization + complexity
  - *Minimum Significant Approximant Bound* (MSAB)
    — complexity (+ characterization)
  - *Expectation Maximization-like algorithm*
    — 2 heuristics + experimental results

The CBA-λ problem

CBA wrt d_λ

Instance: An MC M, and $k \in \mathbb{N}$

Output: An MC Ñ, with at most k states minimizing $d_\lambda(m_0, \tilde{n}_0)$

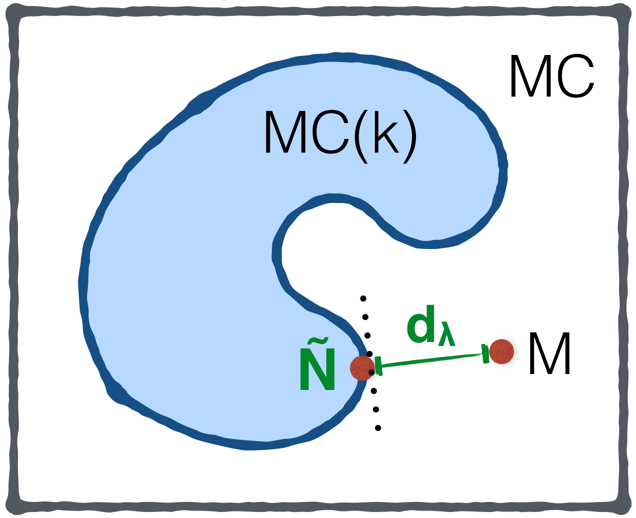


The CBA-λ problem

CBA wrt d_λ

Instance: An MC M, and $k \in \mathbb{N}$

Output: An MC Ñ, with at most k states minimizing $d_\lambda(m_0, \tilde{n}_0)$



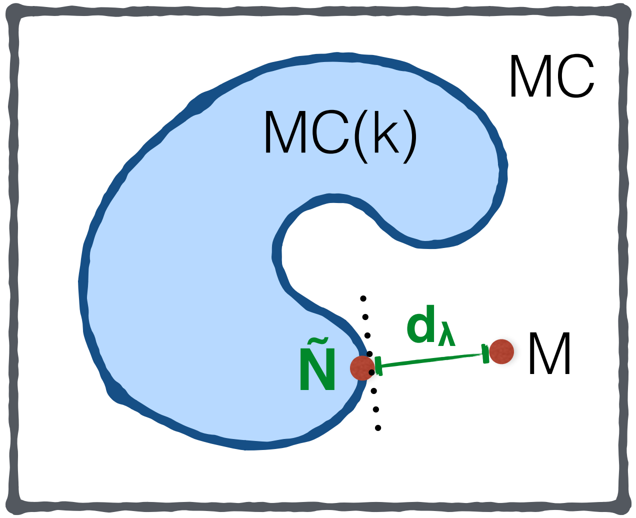
$$d_\lambda(m_0, \tilde{n}_0) = \inf \{ d_\lambda(m_0, n_0) \mid N \in MC(k) \}$$

The CBA-λ problem

CBA wrt d_λ

Instance: An MC M, and $k \in \mathbb{N}$

Output: An MC $\tilde{N}$, with at most k states minimizing $d_\lambda(m_0, \tilde{n}_0)$



$$d_\lambda(m_0, \tilde{n}_0) = \inf \{ d_\lambda(m_0, n_0) \mid N \in MC(k) \}$$

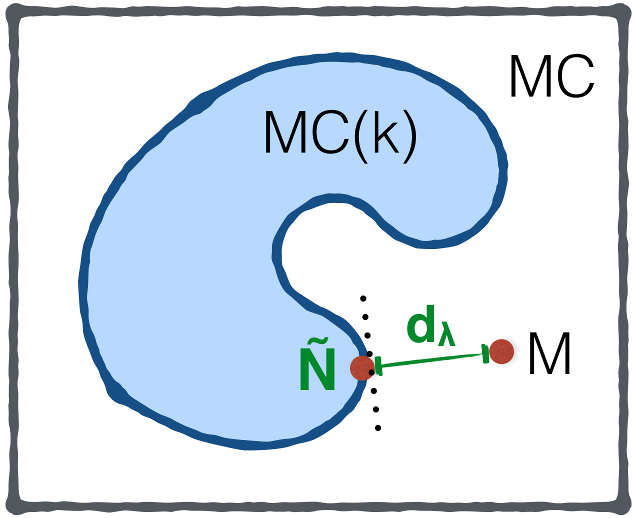
we get a solution iff the infimum is a minimum

The CBA-λ problem

CBA wrt d_λ

Instance: An MC M, and k∈ℕ

Output: An MC Ñ, with at most k states minimizing $d_\lambda(m_0,\tilde{n}_0)$



$$d_\lambda(m_0,\tilde{n}_0) = \inf \{ d_\lambda(m_0,n_0) \mid N \in MC(k) \}$$

we get a solution iff the infimum is a minimum

generalization of bisimilarity quotient

CBA-λ as a Bilinear Program

$d_\lambda(m_0, \tilde{n}_0) = \inf \{ d_\lambda(m_0, n_0) \mid N \in MC(k) \}$

CBA-λ as a Bilinear Program

$$d_\lambda(m_0,\tilde{n}_0) = \inf \{ d_\lambda(m_0,n_0) \mid N \in MC(k) \}$$
$$= \inf \{ d(m_0,n_0) \mid \Gamma_\lambda(d) \leq d,\ N \in MC(k) \}$$

CBA-λ as a Bilinear Program

$d_\lambda(m_0,\tilde{n}_0) = \inf \{ d_\lambda(m_0,n_0) \mid N \in MC(k) \}$
$= \inf \{ d(m_0,n_0) \mid \Gamma_\lambda(d) \leq d,\ N \in MC(k) \}$

$$
\begin{array}{lll}
\text{mimimize} & d_{m_0,n_0} & \\
\text{such that} & d_{m,n} = 1 & \ell(m) \neq \alpha(n) \\
& \lambda \sum_{(u,v)\in M\times N} c^{m,n}_{u,v} \cdot d_{u,v} \leq d_{m,n} & \ell(m) = \alpha(n) \\
& \sum_{v\in N} c^{m,n}_{u,v} = \tau(m)(u) & m,u \in M,\ n \in N \\
& \sum_{u\in M} c^{m,n}_{u,v} = \theta_{n,v} & m \in M,\ n,v \in N \\
& c^{m,n}_{u,v} \geq 0 & m,u \in M,\ n,v \in N
\end{array}
$$

CBA-λ as a Bilinear Program

$d_\lambda(m_0, \tilde{n}_0) = \inf \{ d_\lambda(m_0, n_0) \mid N \in MC(k) \}$

$= \inf \{ d(m_0, n_0) \mid \Gamma_\lambda(d) \leq d, N \in MC(k) \}$

mimimize	d_{m_0,n_0}	
such that	$d_{m,n} = 1$	$\ell(m) \neq \alpha(n)$
	$\lambda \sum_{(u,v) \in M \times N} c^{m,n}_{u,v} \cdot d_{u,v} \leq d_{m,n}$	$\ell(m) = \alpha(n)$
	$\sum_{v \in N} c^{m,n}_{u,v} = \tau(m)(u)$	$m, u \in M,\ n \in N$
	$\sum_{u \in M} c^{m,n}_{u,v} = \theta_{n,v}$	$m \in M,\ n, v \in N$
	$c^{m,n}_{u,v} \geq 0$	$m, u \in M,\ n, v \in N$

CBA-λ as a Bilinear Program

$$d_\lambda(m_0,\tilde{n}_0) = \inf \{\, d_\lambda(m_0,n_0) \mid N \in MC(k) \,\}$$
$$= \inf \{\, d(m_0,n_0) \mid \Gamma_\lambda(d) \leq d,\ N \in MC(k) \,\}$$

mimimize	d_{m_0,n_0}	
such that	$d_{m,n} = 1$	$\ell(m) \neq \alpha(n)$
	$\lambda \sum_{(u,v) \in M \times N} c^{m,n}_{u,v} \cdot d_{u,v} \leq d_{m,n}$	$\ell(m) = \alpha(n)$
	$\sum_{v \in N} c^{m,n}_{u,v} = \tau(m)(u)$	$m, u \in M,\ n \in N$
	$\sum_{u \in M} c^{m,n}_{u,v} = \theta_{n,v}$	$m \in M,\ n, v \in N$
	$c^{m,n}_{u,v} \geq 0$	$m,$

variable!

N = {1…k}

CBA-λ as a Bilinear Program

$$d_\lambda(m_0,\tilde{n}_0) = \inf \{\, d_\lambda(m_0,n_0) \mid N \in MC(k) \,\}$$
$$= \inf \{\, d(m_0,n_0) \mid \Gamma_\lambda(d) \leq d,\ N \in MC(k) \,\}$$

$$\begin{array}{lll}
\text{mimimize} & d_{m_0,n_0} & \\
\text{such that} & d_{m,n} = 1 & \\
 & \lambda \sum_{(u,v)\in M\times N} c^{m,n}_{u,v} \cdot d_{u,v} \leq d_{m,n} & \ell(m) = \alpha(n) \\
 & \sum_{v\in N} c^{m,n}_{u,v} = \tau(m)(u) & m,u \in M,\ n \in N \\
 & \sum_{u\in M} c^{m,n}_{u,v} = \theta_{n,v} & m \in M,\ n,v \in N \\
 & c^{m,n}_{u,v} \geq 0 & m,
\end{array}$$

what labels should the MC N have?

variable!

N = {1…k}

CBA-λ as a Bilinear Program

Lemma (Meaningful labels)

For any $N \in MC(k)$, there exists $N' \in MC(k)$ with labels taken from M, such that $d_\lambda(M,N) \geq d_\lambda(M,N')$

CBA-λ as a Bilinear Program

Lemma (Meaningful labels)

For any N∈MC(k), there exists N'∈MC(k) with labels taken from M, such that $d_\lambda(M,N) \geq d_\lambda(M,N')$

$$
\begin{array}{lll}
\text{mimimize} & d_{m_0,n_0} & \\
\text{such that} & \lambda \sum_{(u,v)\in M\times N} c^{m,n}_{u,v} \cdot d_{u,v} \leq d_{m,n} & m \in M,\, n \in N \\
& 1 - \alpha_{n,l} \leq d_{m,n} \leq 1 & n \in N,\, l \in L(\mathcal{M}),\, \ell(m) \neq l \\
& \alpha_{n,l} \cdot \alpha_{n,l'} = 0 & n \in N,\, l, l' \in L(\mathcal{M}),\, l \neq l' \\
& \sum_{l \in L(\mathcal{M})} \alpha_{n,l} = 1 & n \in N \\
& \sum_{v\in N} c^{m,n}_{u,v} = \tau(m)(u) & m, u \in M,\, n \in N \\
& \sum_{u\in M} c^{m,n}_{u,v} = \theta_{n,v} & m \in M,\, n, v \in N \\
& c^{m,n}_{u,v} \geq 0 & m, u \in M,\, n, v \in N
\end{array}
$$

CBA-λ as a Bilinear Program

Lemma (Meaningful labels)

For any N∈MC(k), there exists N'∈MC(k) with labels taken from M, such that $d_\lambda(M,N) \geq d_\lambda(M,N')$

$$
\begin{array}{lll}
\text{mimimize} & d_{m_0,n_0} & \\
\text{such that} & \lambda \sum_{(u,v)\in M\times N} c^{m,n}_{u,v} \cdot d_{u,v} \leq d_{m,n} & m \in M,\ n \in N \\
& 1 - \alpha_{n,l} \leq d_{m,n} \leq 1 & n \in N,\ l \in L(\mathcal{M}),\ \ell(m) \neq l \\
& \alpha_{n,l} \cdot \alpha_{n,l'} = 0 & n \in N,\ l, l' \in L(\mathcal{M}),\ l \neq l' \\
& \sum_{l \in L(\mathcal{M})} \alpha_{n,l} = 1 & n \in N \\
& \sum_{v \in N} c^{m,n}_{u,v} = \tau(m)(u) & m, u \in M,\ n \in N \\
& \sum_{u \in M} c^{m,n}_{u,v} = \theta_{n,v} & m \in M,\ n, v \in N \\
& c^{m,n}_{u,v} \geq 0 & m, u \in M,\ n, v \in N
\end{array}
$$

CBA-λ as a Bilinear Program

this characterization has two main consequences…

1. CBA-λ admits always a solution (finite intersection of closed subsets)
2. CBA-λ can be approximated up to any precision

Complexity of CBA-λ

actually, its decision variant!

Complexity of CBA-λ

actually, its decision variant!

Complexity Upper-bound

BA-λ is in **PSPACE**

Proof sketch: *we can encode the question ⟨M,k,ε⟩∈BA-λ to that of checking the feasibility of a set of bilinear inequalities. This can be encoded as a decision problem for the existential theory of the reals, thus it can be solved in PSPACE [Canny—STOC88].*

Complexity of CBA-λ

actually, its decision variant!

Complexity Upper-bound

BA-λ is in **PSPACE**

Complexity lower-bound

BA-λ is **NP-hard**

Proof idea: *we provide a reduction from VERTEX COVER.*
(see the appendix for a sketch of the reduction)

Complexity of CBA-λ

actually, its decision variant!

Complexity Upper-bound

BA-λ is in **PSPACE**

Complexity lower-bound

BA-λ is **NP-hard**

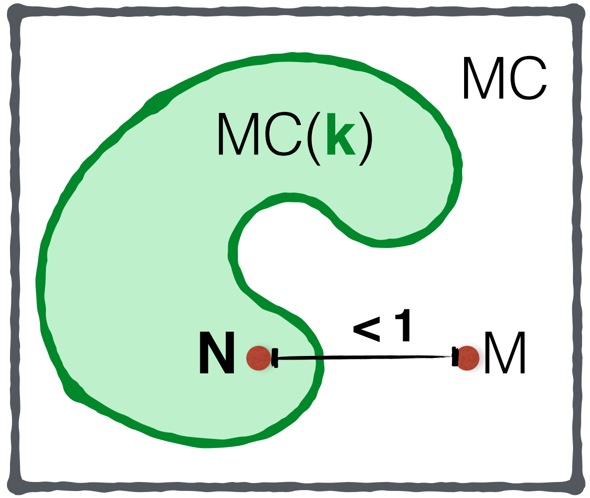
unlikely to solve CBA as simple linear program

The MSAB-λ problem

The MSAB wrt d_λ

Instance: An MC M

Output: The smallest k such that $d_\lambda(m_0,n_0)<1$, for some $N \in MC(k)$

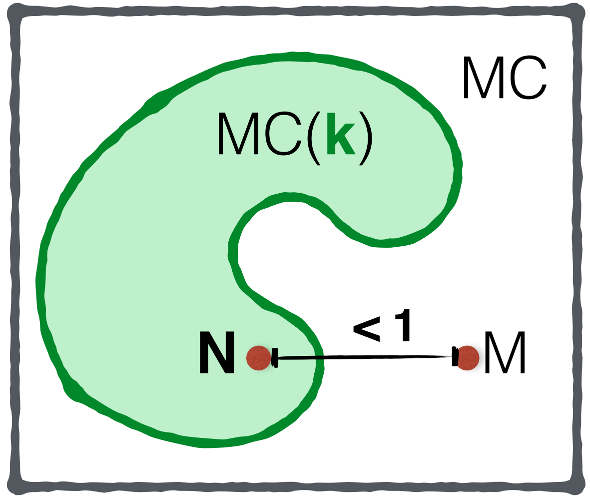


The MSAB-λ problem

The MSAB wrt d_λ

Instance: An MC M

Output: The smallest k such that $d_\lambda(m_0,n_0)<1$, for some N∈MC(k)



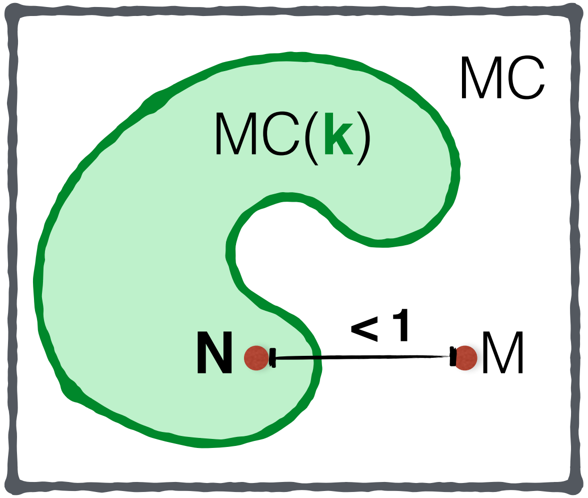
For **λ<1**, the MSAB-λ problem is trivial, because the solution is always k=1

The MSAB-λ problem

The MSAB wrt d_λ

Instance: An MC M

Output: The smallest k such that $d_\lambda(m_0,n_0)<1$, for some N∈MC(k)



For **λ<1**, the MSAB-λ problem is trivial, because the solution is always k=1

For **λ=1**, the same problem is surprisingly difficult…

Complexity of MSAB-1

actually, its decision variant!

Theorem

SBA-1 is **NP-complete**

Proof idea: *we provide a reduction from VERTEX COVER.*
(see the appendix for a sketch of the reduction)

Towards an Algorithm…

Towards an Algorithm…

- The CBA can be solved as a bilinear program. Theoretically nice, but practically unfeasible!

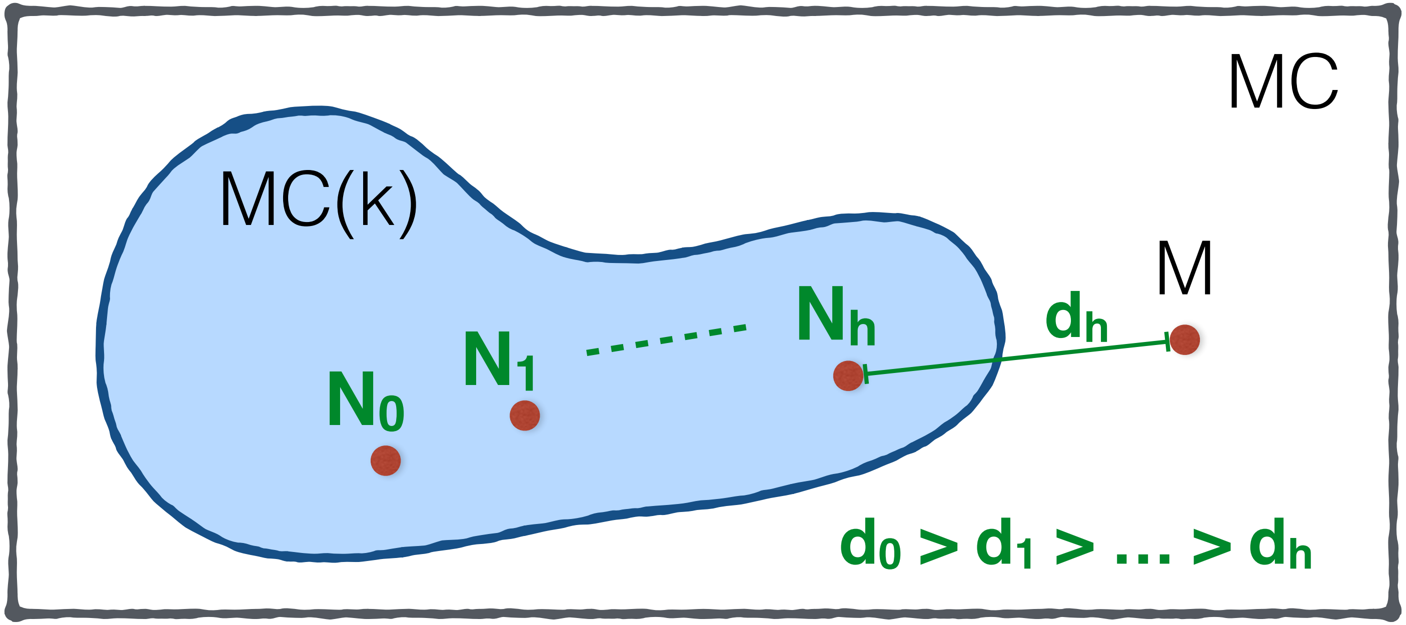
  (our implementation in PENBMI can handle MCs with at most 5 states…)

Towards an Algorithm…

- The CBA can be solved as a bilinear program. Theoretically nice, but practically unfeasible!
  (our implementation in PENBMI can handle MCs with at most 5 states…)

- We are happy with **sub-optimal solutions** if they can be obtained by a practical algorithm.

EM-like Algorithm



- Given the MC M and an initial approximant N_0
- it produces a sequence $N_0, \ldots, N_h$ of approximants having strictly decreasing distance from M
- N_h may be a sub-optimal solution of CBA-λ

EM-like Algorithm

Algorithm 1

Input: $\mathcal{M} = (M, \tau, \ell)$, $\mathcal{N}_0 = (N, \theta_0, \alpha)$, and $h \in \mathbb{N}$.

1. $i \leftarrow 0$
2. **repeat**
3. $\quad i \leftarrow i + 1$
4. $\quad$ compute $\mathcal{C} \in \Omega(\mathcal{M}, \mathcal{N}_{i-1})$ such that $\delta_\lambda(\mathcal{M}, \mathcal{N}_{i-1}) = \gamma_\lambda^{\mathcal{C}}(\mathcal{M}, \mathcal{N}_{i-1})$
5. $\quad \theta_i \leftarrow \textsc{UpdateTransition}(\theta_{i-1}, \mathcal{C})$
6. $\quad \mathcal{N}_i \leftarrow (N, \theta_i, \alpha)$
7. **until** $\delta_\lambda(\mathcal{M}, \mathcal{N}_i) > \delta_\lambda(\mathcal{M}, \mathcal{N}_{i-1})$ or $i \geq h$
8. **return** $\mathcal{N}_{i-1}$

EM-like Algorithm

Algorithm 1

Input: $\mathcal{M} = (M, \tau, \ell)$, $\mathcal{N}_0 = (N, \theta_0, \alpha)$, and $h \in \mathbb{N}$.

1. $i \leftarrow 0$
2. **repeat**
3. $\quad i \leftarrow i + 1$
4. $\quad$ compute $\mathcal{C} \in \Omega(\mathcal{M}, \mathcal{N}_{i-1})$ such that $\delta_\lambda(\mathcal{M}, \mathcal{N}_{i-1}) = \gamma_\lambda^{\mathcal{C}}(\mathcal{M}, \mathcal{N}_{i-1})$
5. $\quad \theta_i \leftarrow \textsc{UpdateTransition}(\theta_{i-1}, \mathcal{C})$
6. $\quad \mathcal{N}_i \leftarrow (N, \theta_i, \alpha)$
7. **until** $\delta_\lambda(\mathcal{M}, \mathcal{N}_i) > \delta_\lambda(\mathcal{M}, \mathcal{N}_{i-1})$ or $i \geq h$
8. **return** $\mathcal{N}_{i-1}$

Intuitive Idea

UpdateTransition assigns greater probability to transitions that are most representative of the behavior of M

Two update heuristics

- **Averaged Marginal (AM)**: given N_k we construct N_{k+1} by averaging the marginal of certain "coupling variables" obtained by optimizing the number of occurrences of the edges that are most likely to be seen in M.

- **Averaged Expectations (AE)**: similar to the above, but now the N_{k+1} looks only the expectation of the number of occurrences of the edges likely to be found in M.

Two update heuristics

- **Averaged Marginal (AM)**: given N_k we construct N_{k+1} by averaging the marginal of certain "coupling variables" obtained by optimizing the number of occurrences of the edges that are most likely to be seen in M.

- **Averaged Expectations (AE)**: similar to the above, but now the N_{k+1} looks only th… of the number of occurren… likely to be found in M.

UpdateTransition in polynomial time for both heuristics!

Case	$\|M\|$	k	$\lambda = 1$				$\lambda = 0.8$			
			δ_λ-init	δ_λ-final	#	time	δ_λ-init	δ_λ-final	#	time
IPv4 (AM)	23	5	0.775	0.054	3	4.8	0.576	0.025	3	4.8
	53	5	0.856	0.062	3	25.7	0.667	0.029	3	25.9
	103	5	0.923	0.067	3	116.3	0.734	0.035	3	116.5
	53	6	0.757	0.030	3	39.4	0.544	0.011	3	39.4
	103	6	0.837	0.032	3	183.7	0.624	0.017	3	182.7
	203	6	–	–	–	TO	–	–	–	TO
IPv4 (AE)	23	5	0.775	0.109	2	2.7	0.576	0.049	3	4.2
	53	5	0.856	0.110	2	14.2	0.667	0.049	3	21.8
	103	5	0.923	0.110	2	67.1	0.734	0.049	3	100.4
	53	6	0.757	0.072	2	21.8	0.544	0.019	3	33.0
	103	6	0.837	0.072	2	105.9	0.624	0.019	3	159.5
	203	6	–	–	–	TO	–	–	–	TO
DrkW (AM)	39	7	0.565	0.466	14	259.3	0.432	0.323	14	252.8
	49	7	0.568	0.460	14	453.7	0.433	0.322	14	420.5
	59	8	0.646	–	–	TO	0.423	–	–	TO
DrkW (AE)	39	7	0.565	0.435	11	156.6	0.432	0.321	2	28.6
	49	7	0.568	0.434	10	247.7	0.433	0.316	2	46.2
	59	8	0.646	0.435	10	588.9	0.423	0.309	2	115.7

Table 1. Comparison of the performance of EM algorithm on the IPv4 zeroconf protocol and the classic Drunkard's Walk w.r.t. the heuristics AM and AE.

Future Work

- **Conjecture 1:** (with Nathanaël Fijalkow)
  Is BA-1 is SUM-OF-SQUARE-ROOTS-hard

- **Conjecture 2:** (by Borja Balle)
  for $\lambda<1$, BA-λ is in NP (hence NP-complete!)

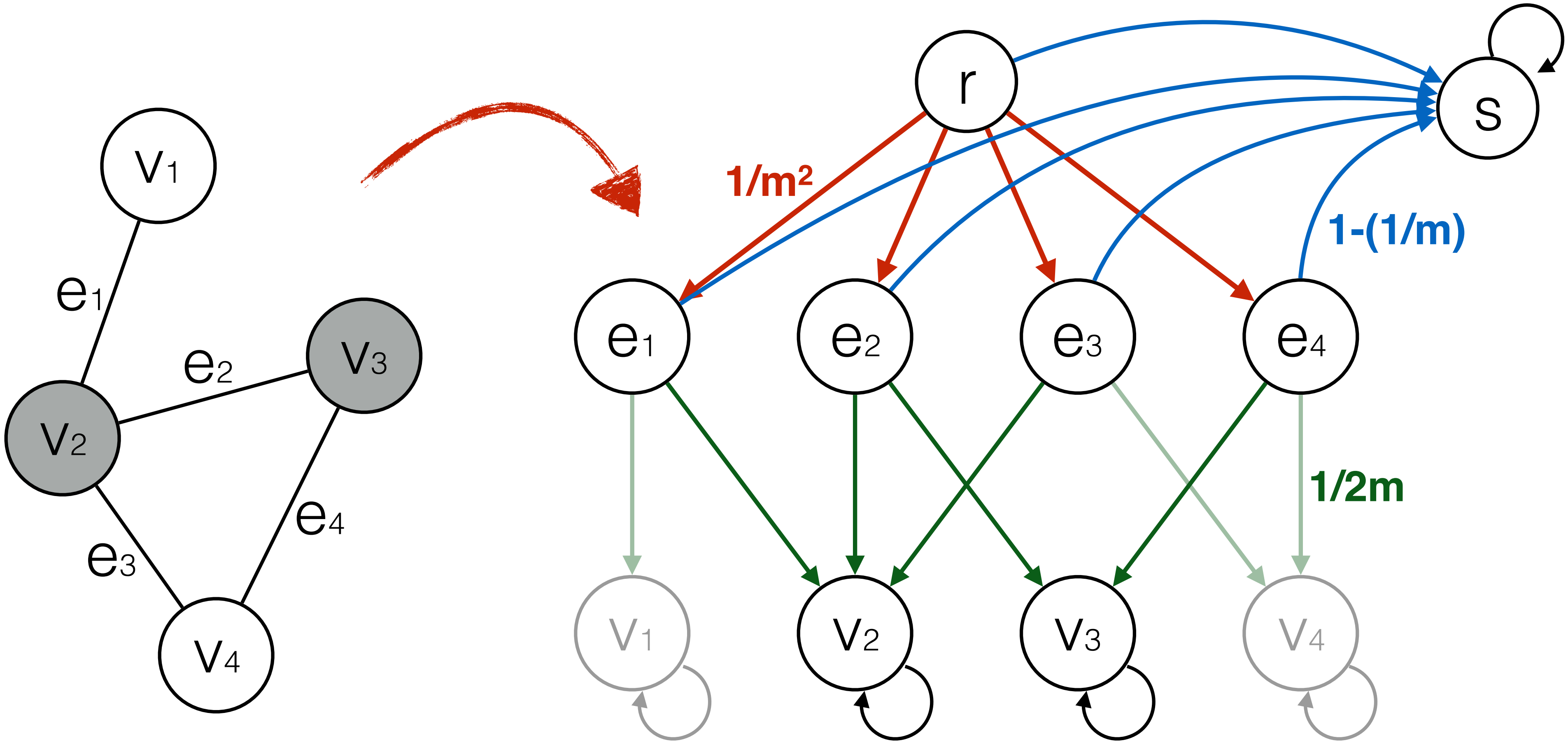
- Real/better EM-heuristics?

- What about different models/distances?

Thank you
for your attention

Appendix

BA-λ is NP-hard

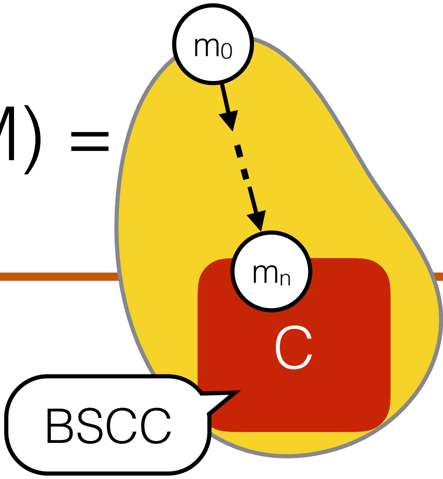


⟨G,h⟩∈VERTEX COVER iff ⟨M_G, m+h+2, $\lambda^2/2m^2$⟩∈BA-λ

Characterization of SBA-1

Lemma

Assume M be maximally collapsed. Then,

$\langle M,k \rangle \in$SBA-1 iff $\mathcal{G}(M)$ =  and $h+|C| \leq k$

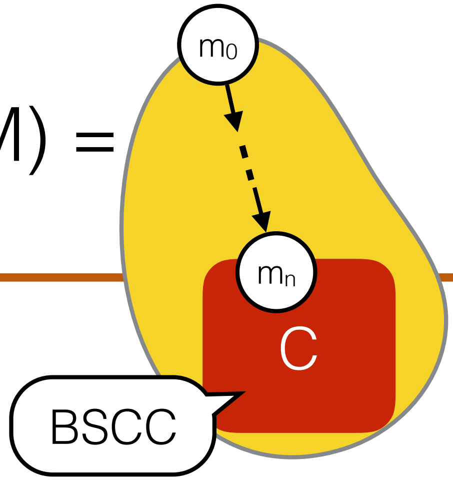
number of labels in $m_0 \ldots m_{n-1}$

Characterization of SBA-1

Lemma

Assume M be maximally collapsed. Then,

$\langle M,k\rangle \in \text{SBA-1}$ iff $\mathcal{G}(M) =$

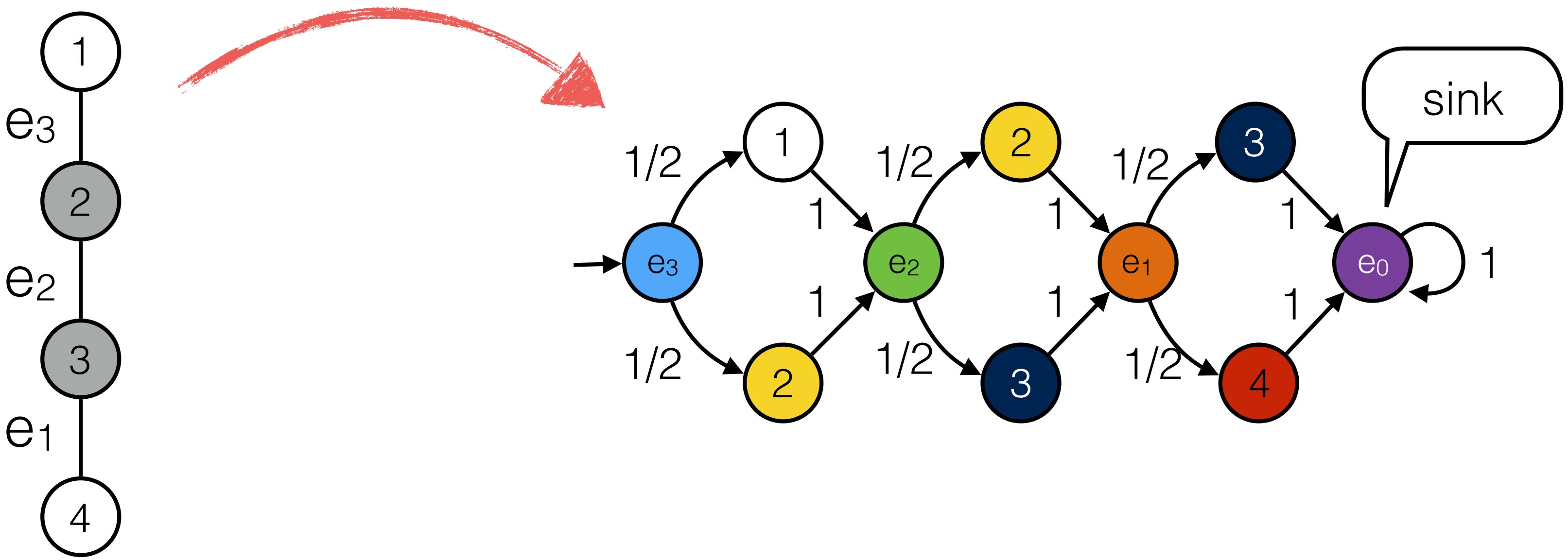


and $h+|C| \leq k$

number of labels in $m_0 \ldots m_{n-1}$

Proof sketch: *compute with Tarjan's algorithm all the SCCs of $\mathcal{G}(M)$. Then non deterministically choose a BSCC and a path to it. In poly-time we can count the number of labels in the path and the size of the BSCC.*

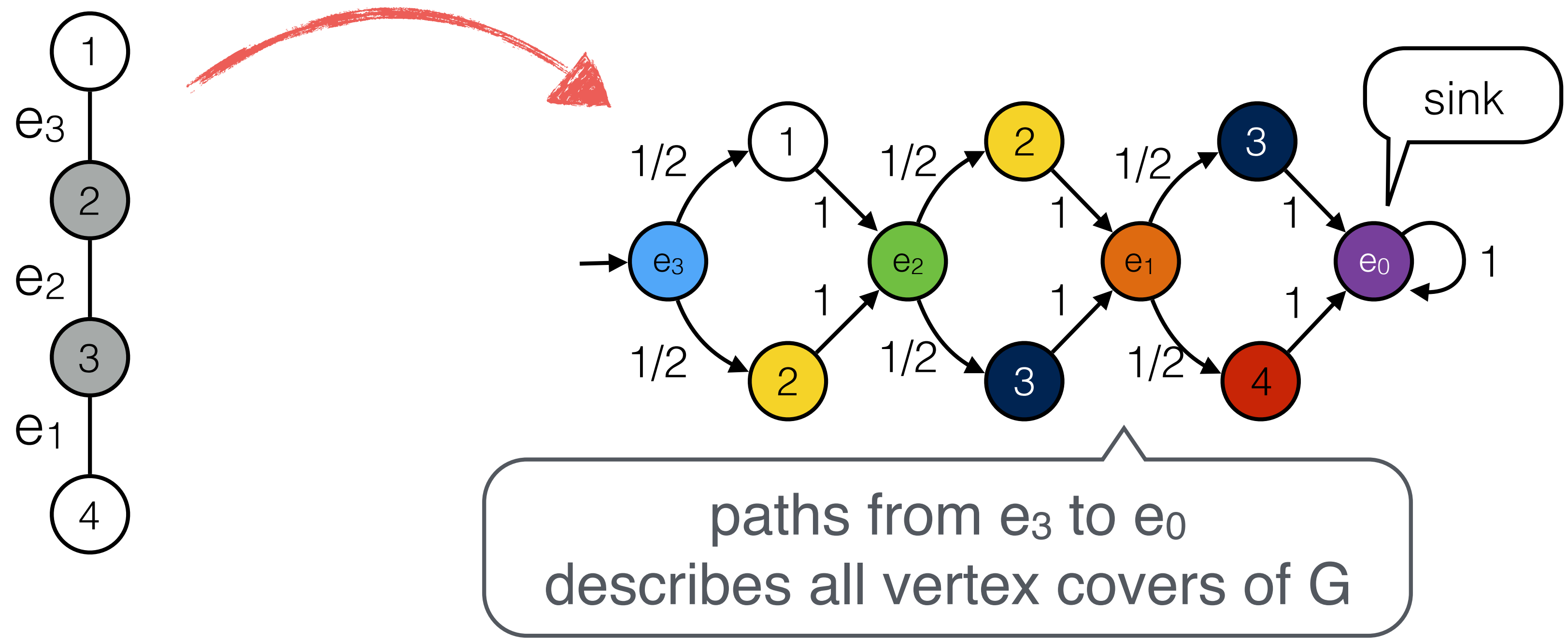
SBA-1 is **NP-hard**



Proof sketch: *by reduction to VERTEX COVER:*

$\langle G,h \rangle \in VERTEX\ COVER$ *iff* $\langle M_G, h+m+1 \rangle \in SBA\text{-}1$

SBA-1 is **NP-hard**



Proof sketch: *by reduction to VERTEX COVER:*

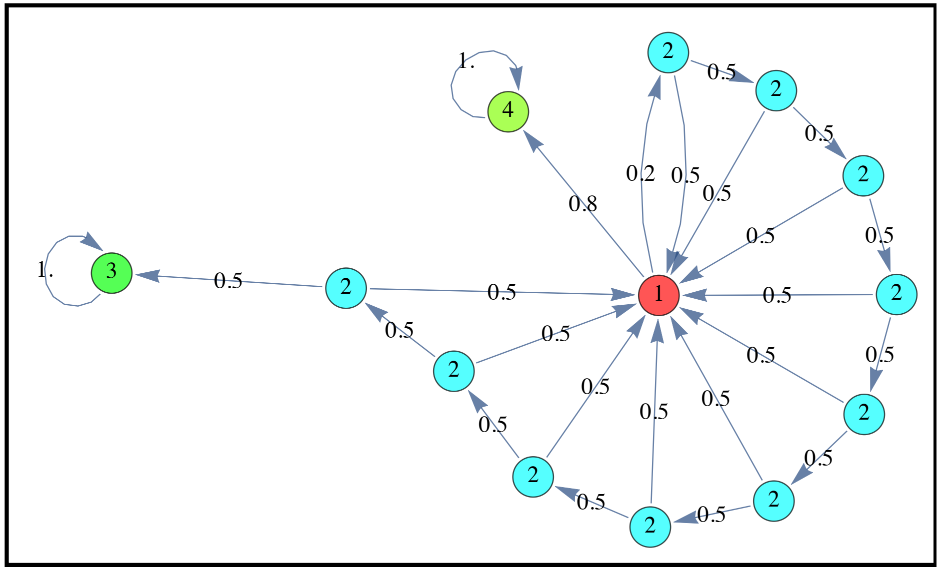
$\langle G,h\rangle \in$ *VERTEX COVER* *iff* $\langle M_G, h+m+1\rangle \in$ *SBA-1*

EM-like algorithm
(experimental results)

IPv4 Zero Conf Protocol

Averaged Marginal (AM)

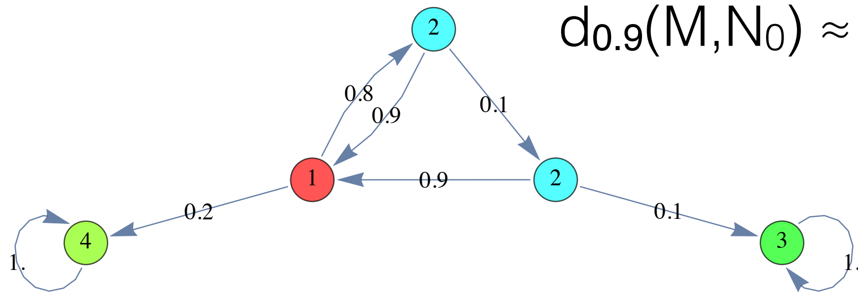
Input model



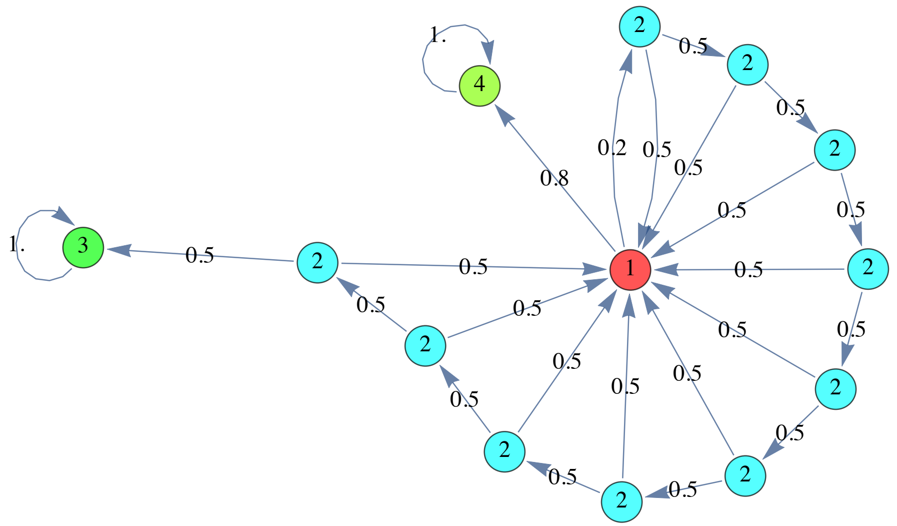
IPv4 Zero Conf Protocol

Averaged Marginal (AM)

$d_{0.9}(M,N_0) \approx 0.67$



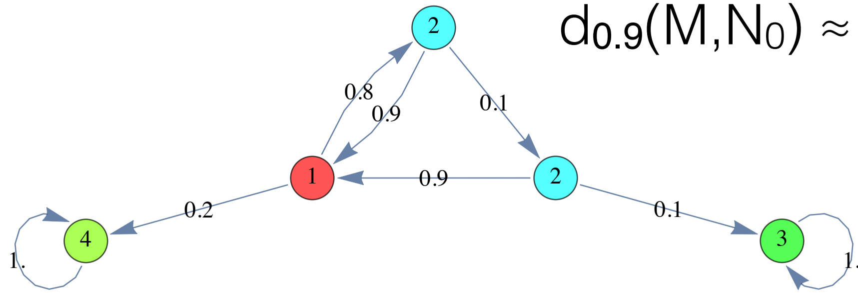
Input model



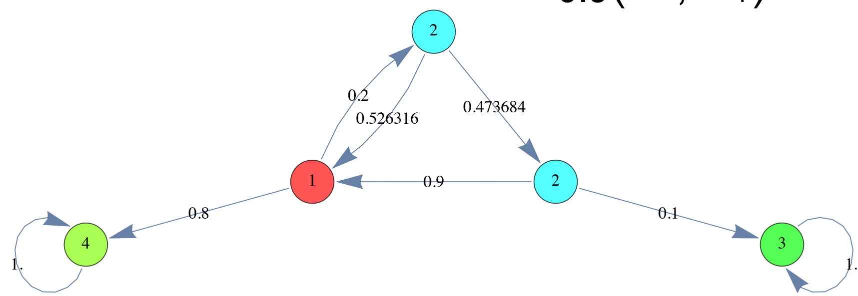
IPv4 Zero Conf Protocol

Averaged Marginal (AM)

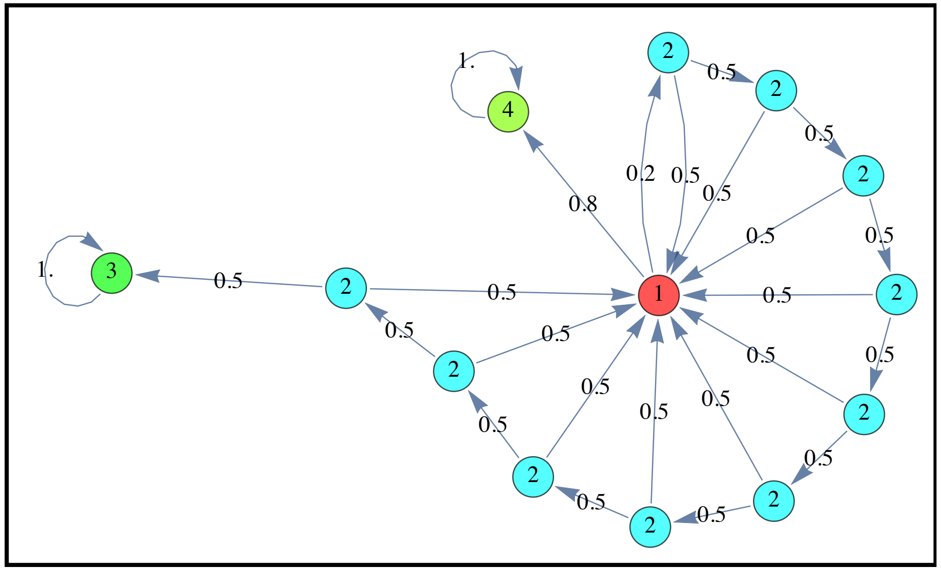
$d_{0.9}(M,N_0) \approx 0.67$



$d_{0.9}(M,N_1) \approx 0.043$



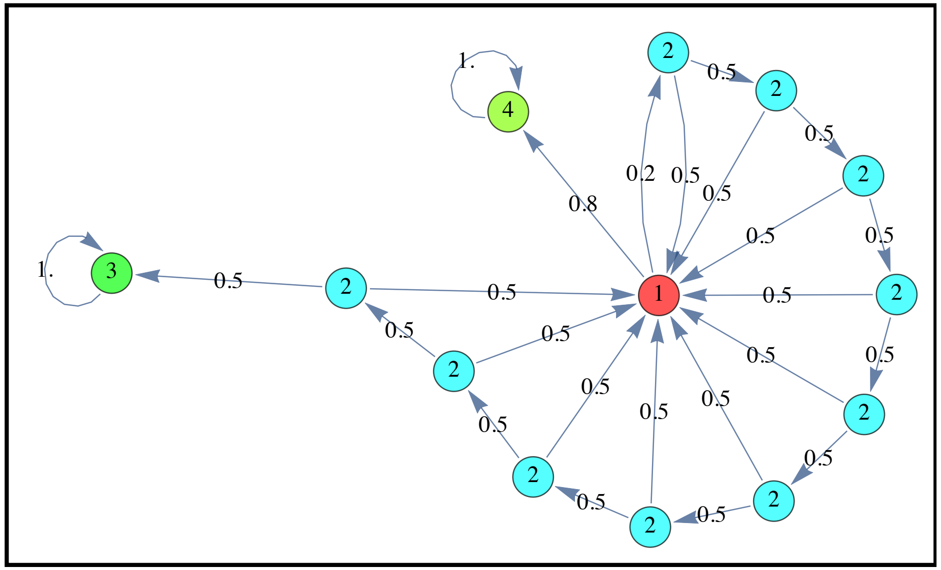
Input model



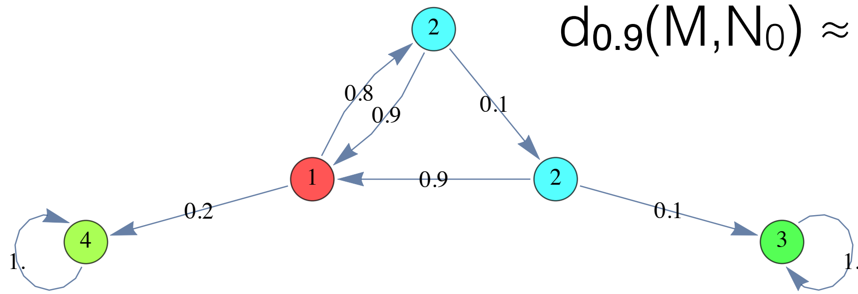
IPv4 Zero Conf Protocol

Averaged Marginal (AM)

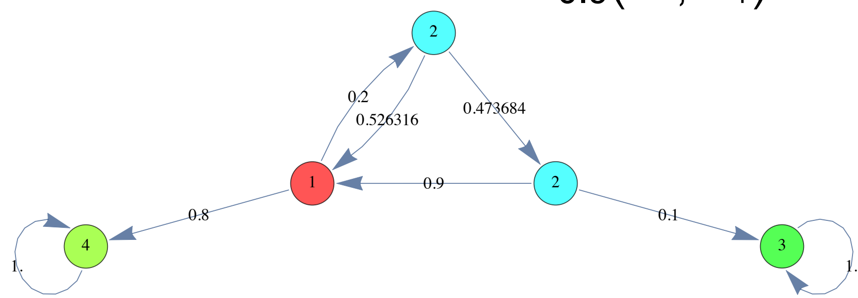
Input model



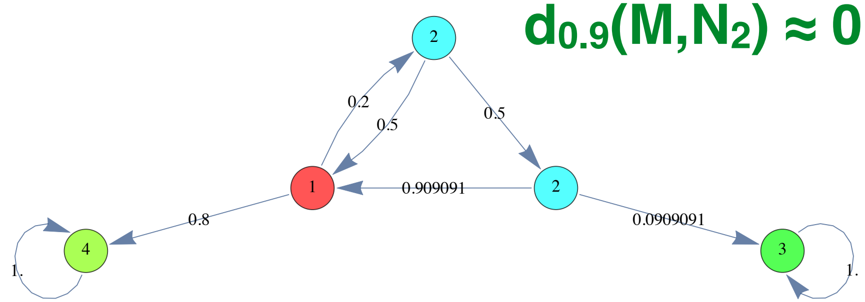
$d_{0.9}(M,N_0) \approx 0.67$



$d_{0.9}(M,N_1) \approx 0.043$



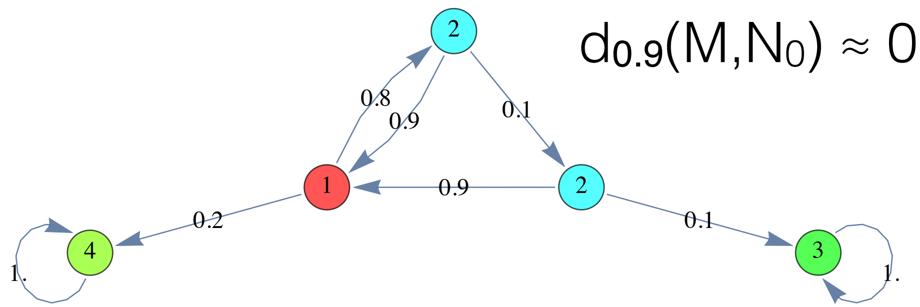
$d_{0.9}(M,N_2) \approx 0.041$



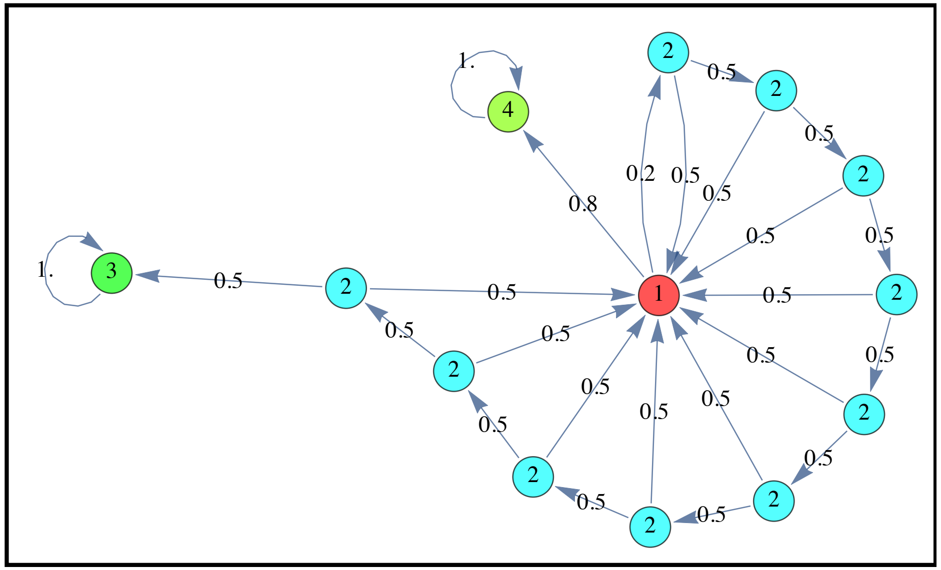
IPv4 Zero Conf Protocol

Averaged Expectations (AE)

$d_{0.9}(M,N_0) \approx 0.67$



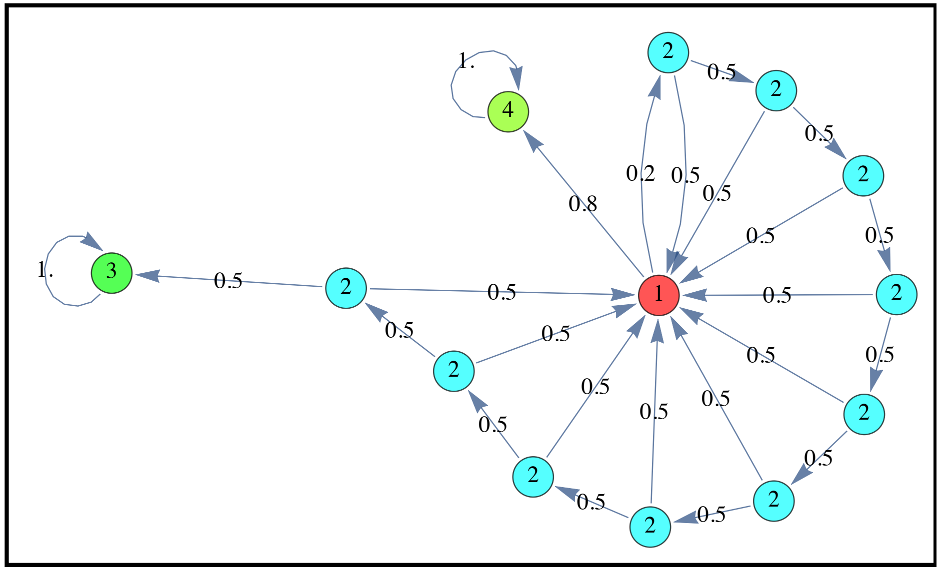
Input model



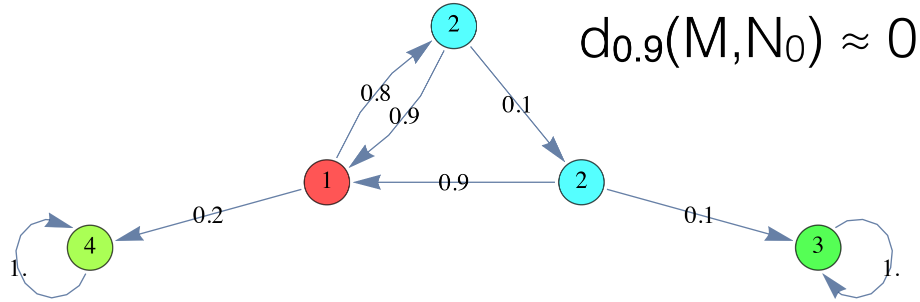
IPv4 Zero Conf Protocol

Averaged Expectations (AE)

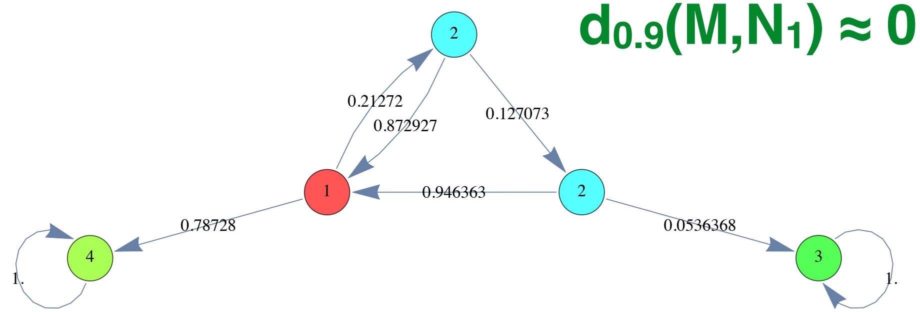
Input model



$d_{0.9}(M,N_0) \approx 0.67$



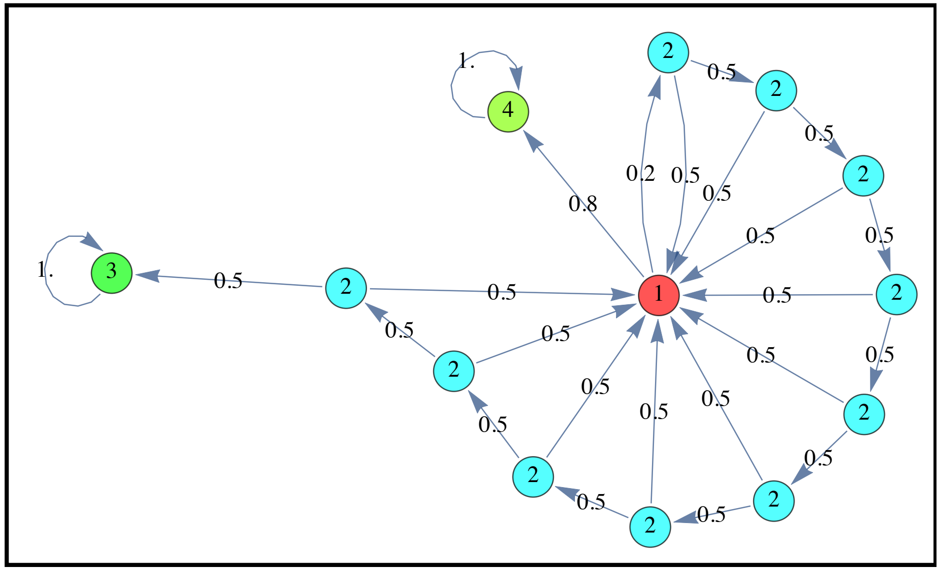
$d_{0.9}(M,N_1) \approx 0.08$



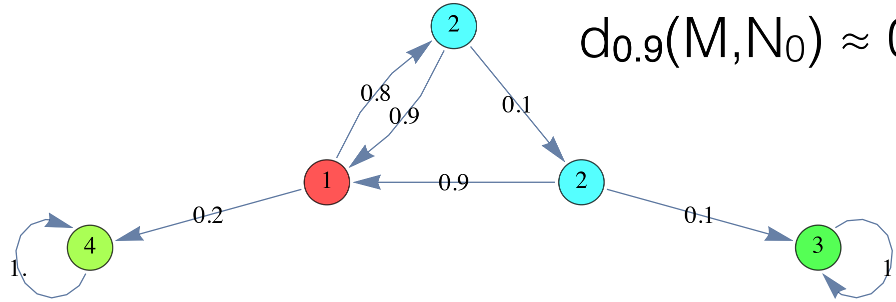
IPv4 Zero Conf Protocol

Averaged Expectations (AE)

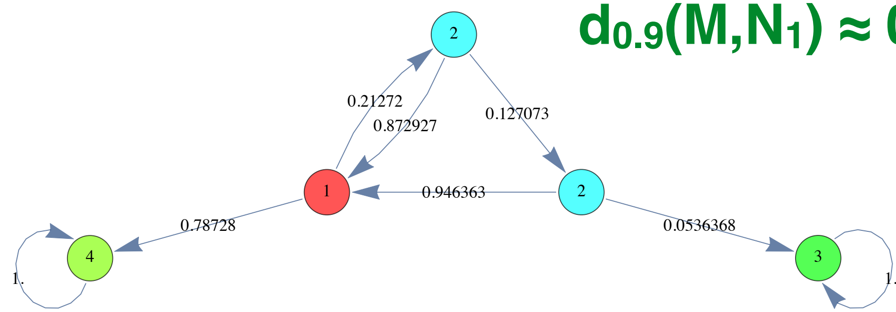
Input model



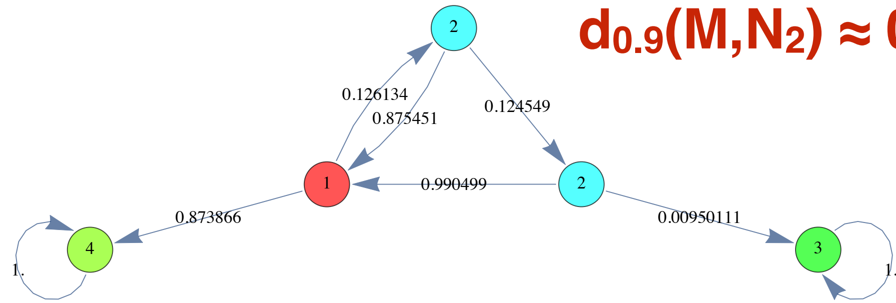
$d_{0.9}(M,N_0) \approx 0.67$



$d_{0.9}(M,N_1) \approx 0.08$



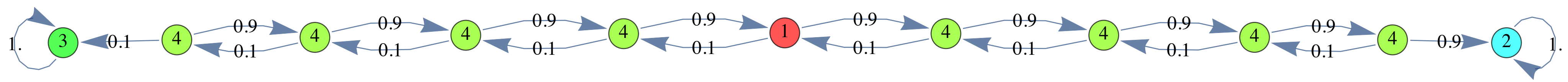
$d_{0.9}(M,N_2) \approx 0.11$



Drunkard's Walk

Averaged Marginal (AM)

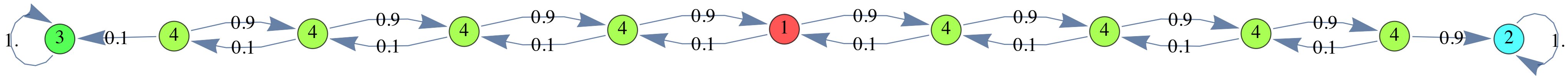
Input model

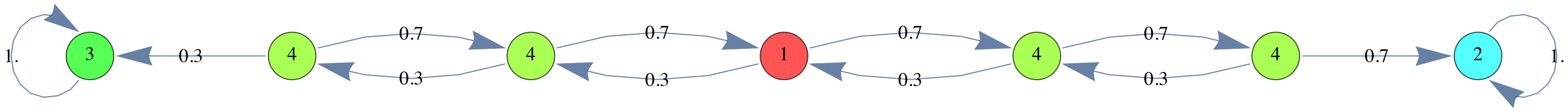


Drunkard's Walk

Averaged Marginal (AM)

Input model



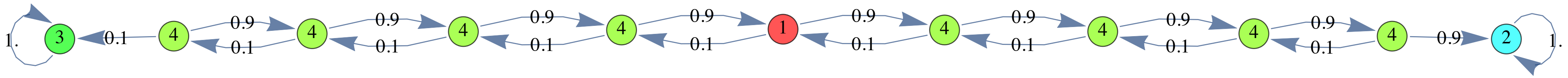


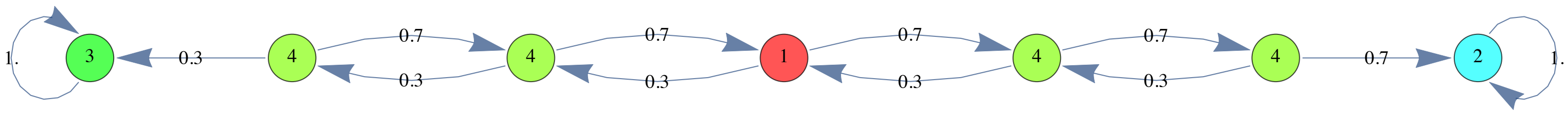
$d_{0.9}(M,N_0) \approx 0.64$

Drunkard's Walk

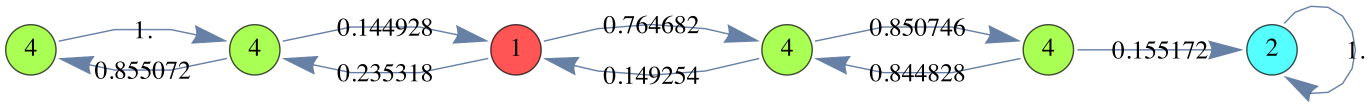
Averaged Marginal (AM)

Input model





$d_{0.9}(M,N_0) \approx 0.64$

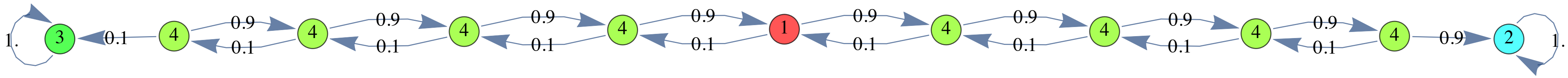


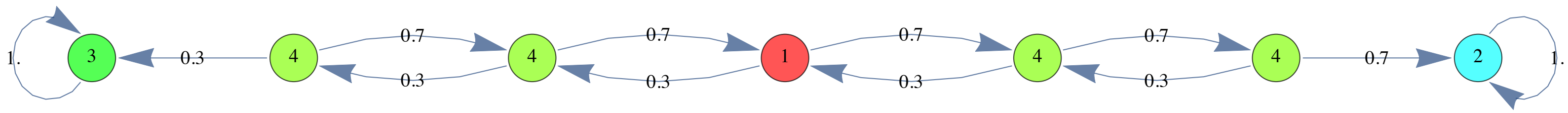
$d_{0.9}(M,N_1) \approx 0.56$

Drunkard's Walk

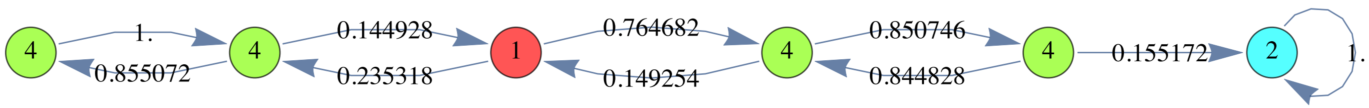
Averaged Marginal (AM)

Input model

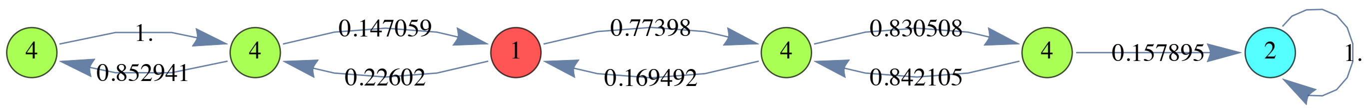




$d_{0.9}(M,N_0) \approx 0.64$



$\mathbf{d_{0.9}(M,N_1) \approx 0.56}$

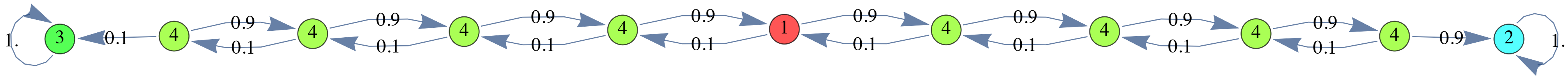


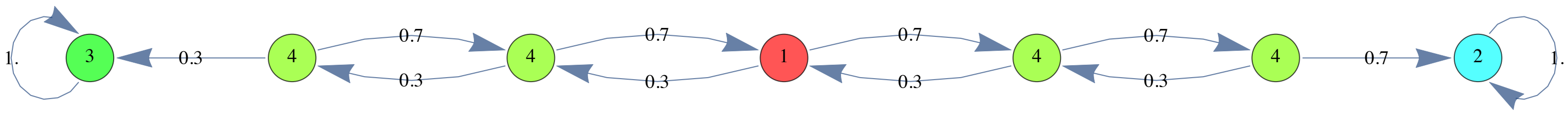
$\mathbf{d_{0.9}(M,N_2) \approx 0.567}$

Drunkard's Walk

Averaged Expectations (AE)

Input model



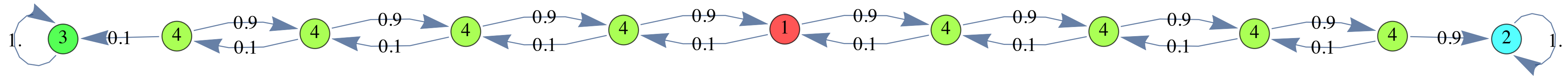


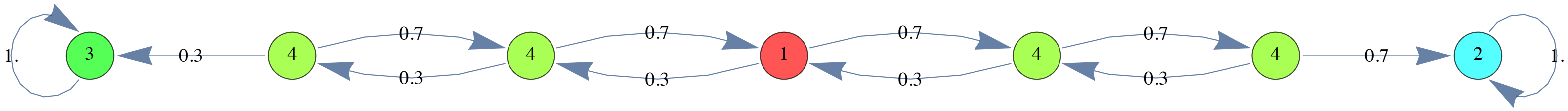
$\delta_{0.9}(M, N_0) \approx 0.64$

Drunkard's Walk

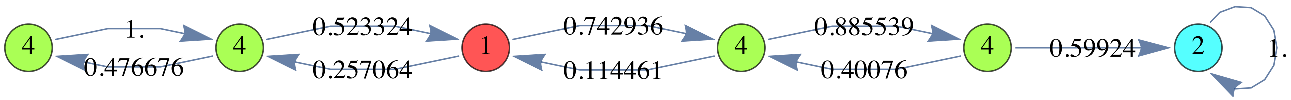
Averaged Expectations (AE)

Input model





$\delta_{0.9}(M,N_0) \approx 0.64$

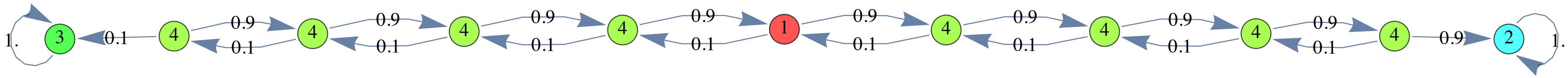


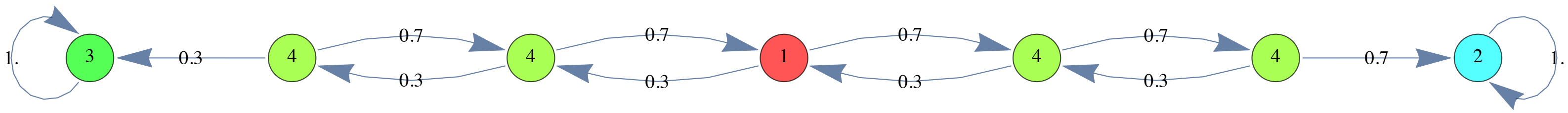
$\delta_{0.9}(M,N_1) \approx 0.56$

Drunkard's Walk

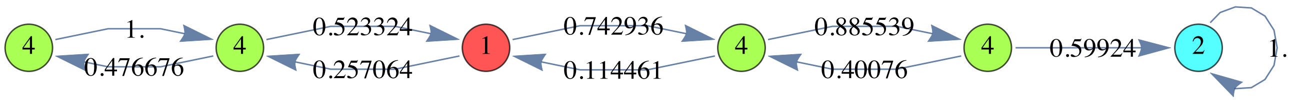
Averaged Expectations (AE)

Input model

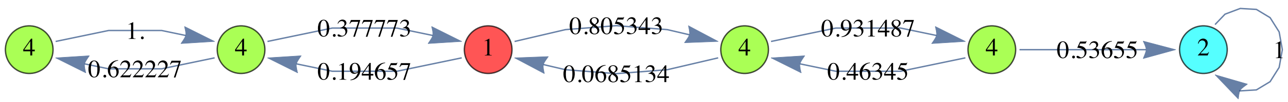




$\delta_{0.9}(M,N_0) \approx 0.64$



$\delta_{0.9}(M,N_1) \approx 0.56$

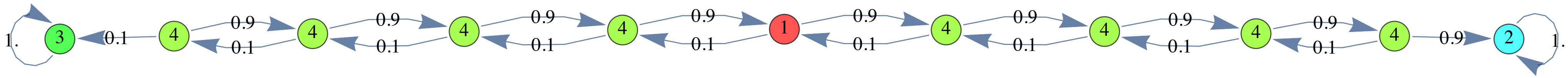


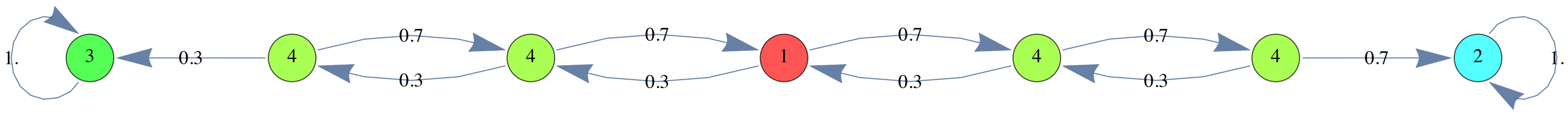
$\delta_{0.9}(M,N_2) \approx 0.543$

Drunkard's Walk

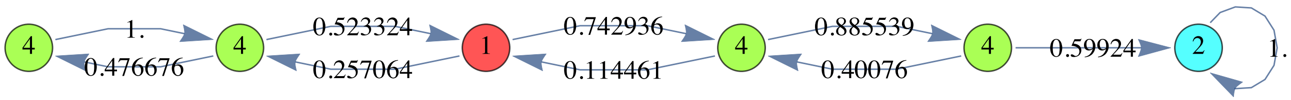
Averaged Expectations (AE)

Input model

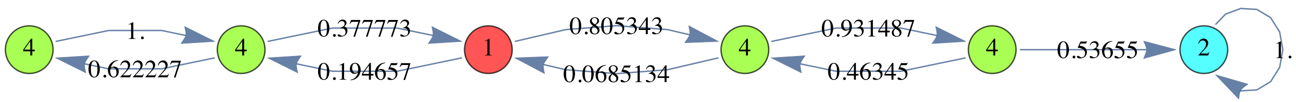




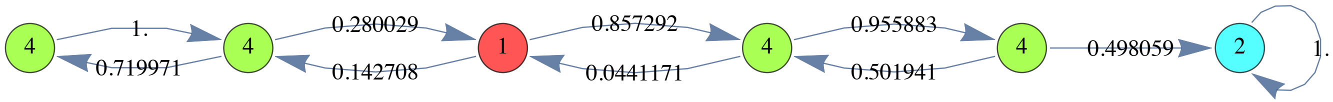
$\delta_{0.9}(M,N_0) \approx 0.64$



$\delta_{0.9}(M,N_1) \approx 0.56$



$\delta_{0.9}(M,N_2) \approx 0.543$



$\boldsymbol{\delta_{0.9}(M,N_3) \approx 0.540}$