Minimizing Markov chains Beyond Bisimilarity

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

1 November 2016 - Rold Storkro, Denmark

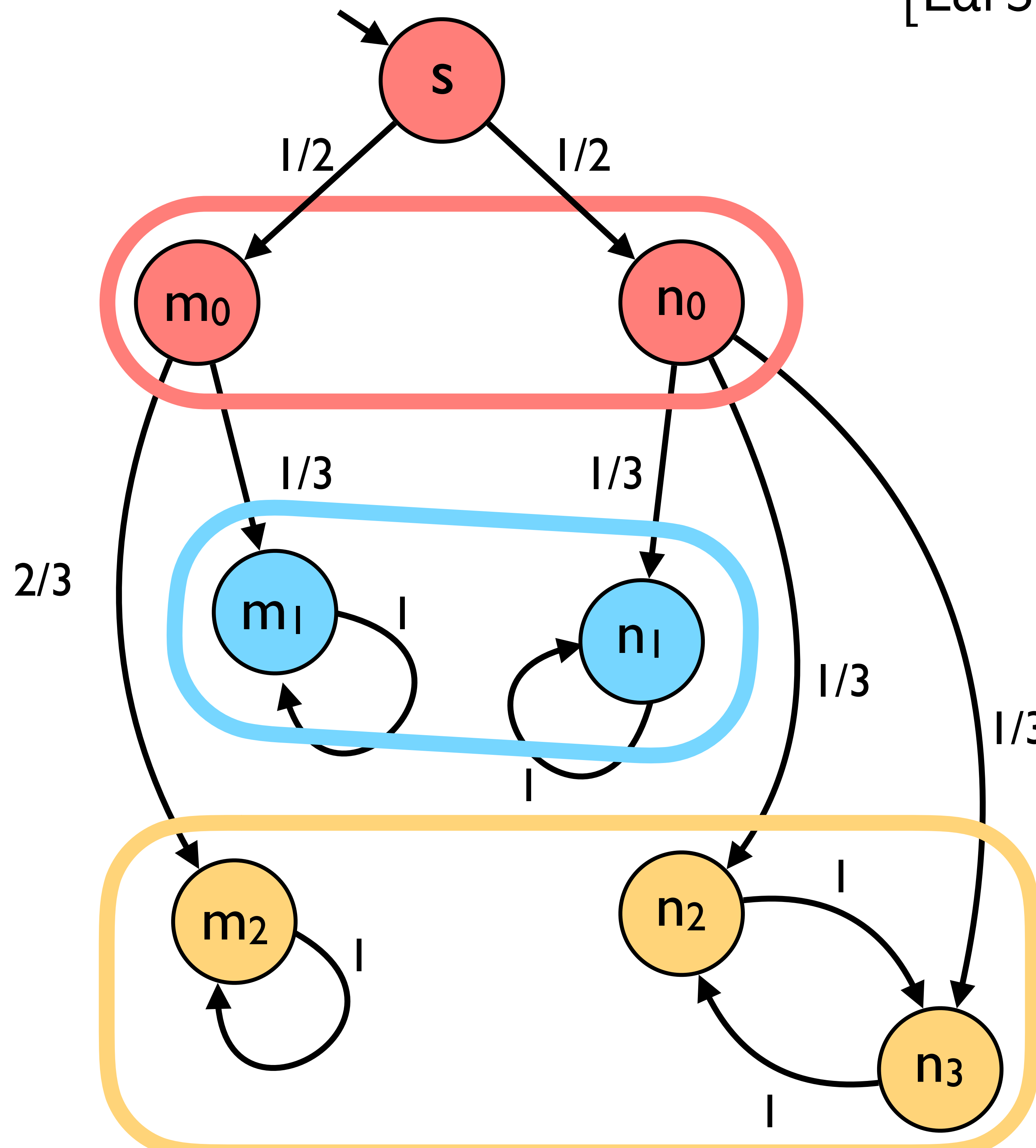
NWPT 2016

The focus of the talk

- Probabilistic Models (Markov chains)
- Automatic verification (e.g., Model Checking)
  - **state space explosion** (even after model reduction, symbolic tech., partial-order reduction)
- Still too large: one needs to compromise in the accuracy of the model (introduce an error)
- Our proposal: **metric-based** state space reduction

Probabilistic Bisimulation

[Larsen & Skou'91]

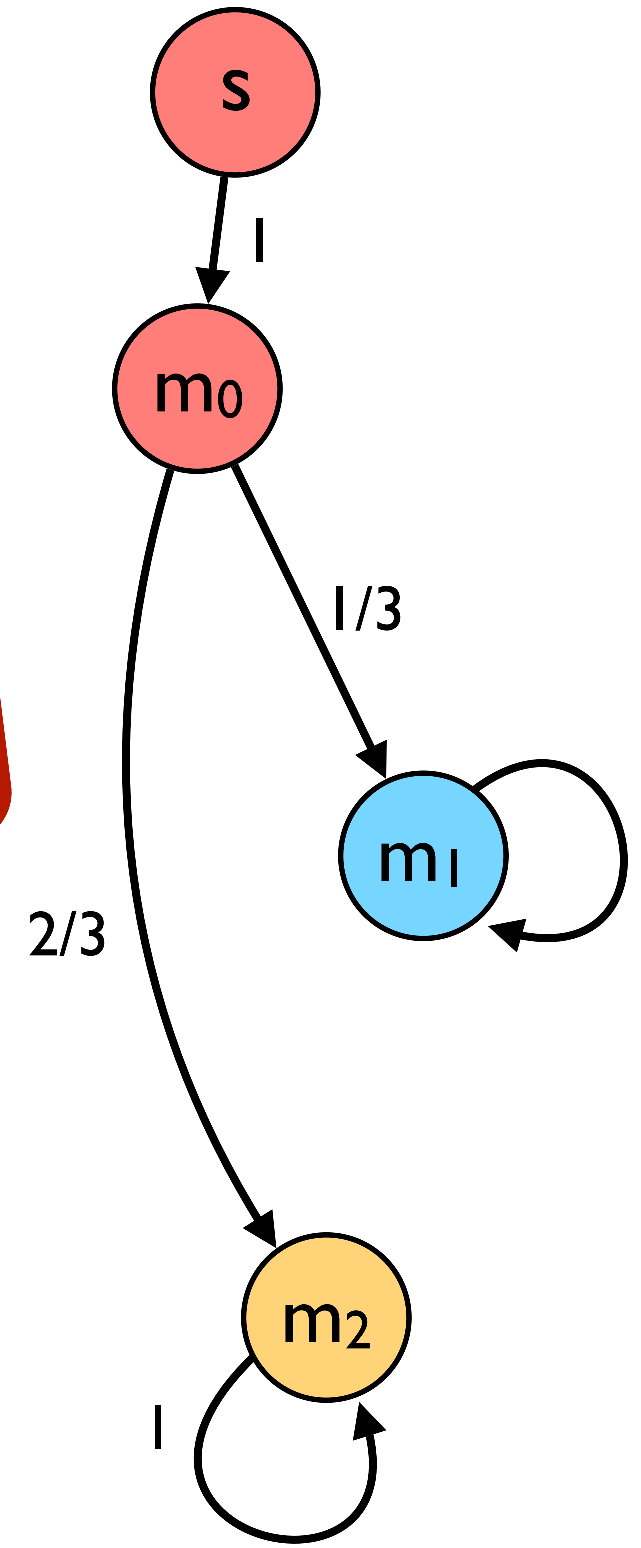


Probabilistic Bisimulation

[Larsen & Skou'91]

Optimal lumping
[Kemeny & Snell'60]

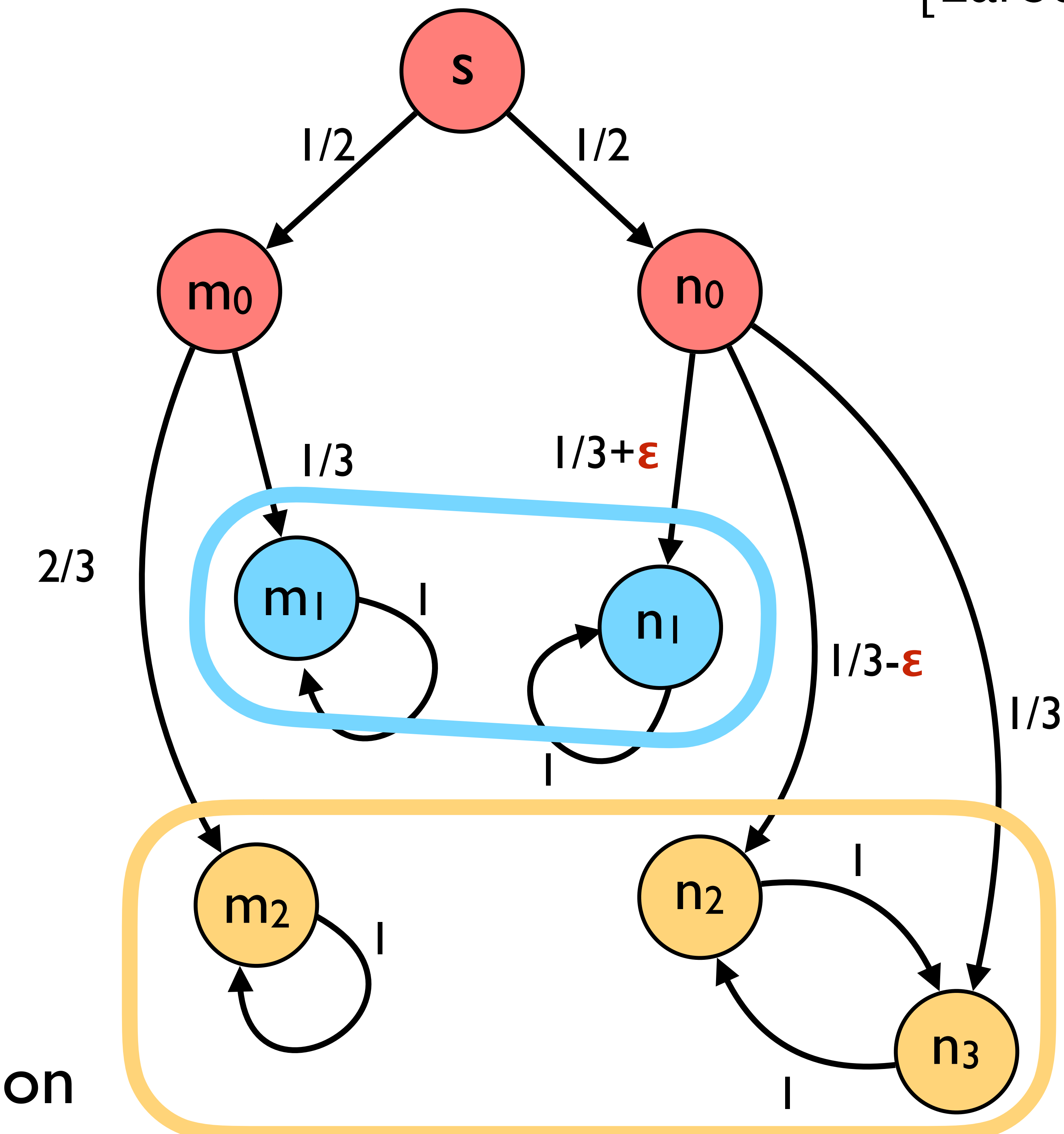
Efficient technique
[Derisavi et al.'03]



...but small variations may prevent aggregation

Probabilistic Bisimulation

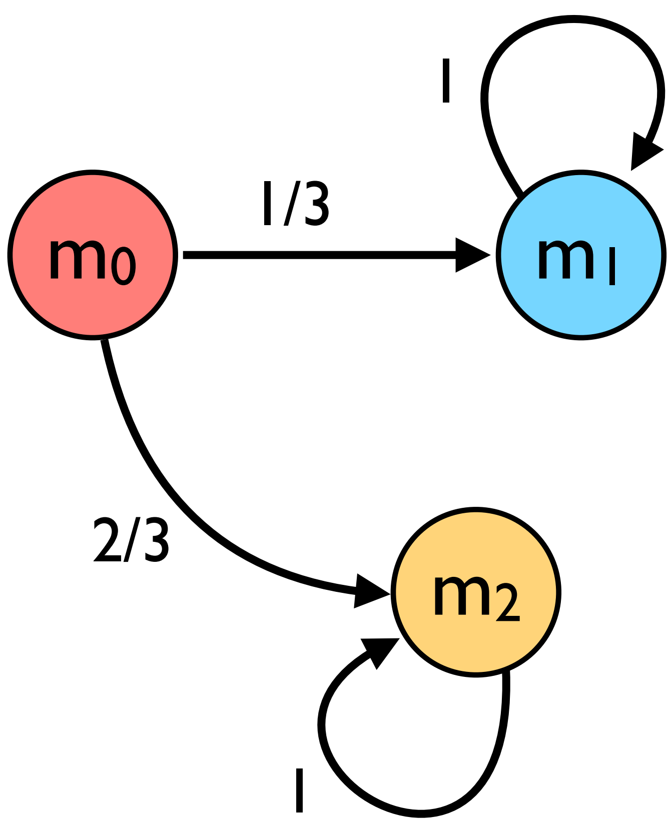
[Larsen & Skou'91]

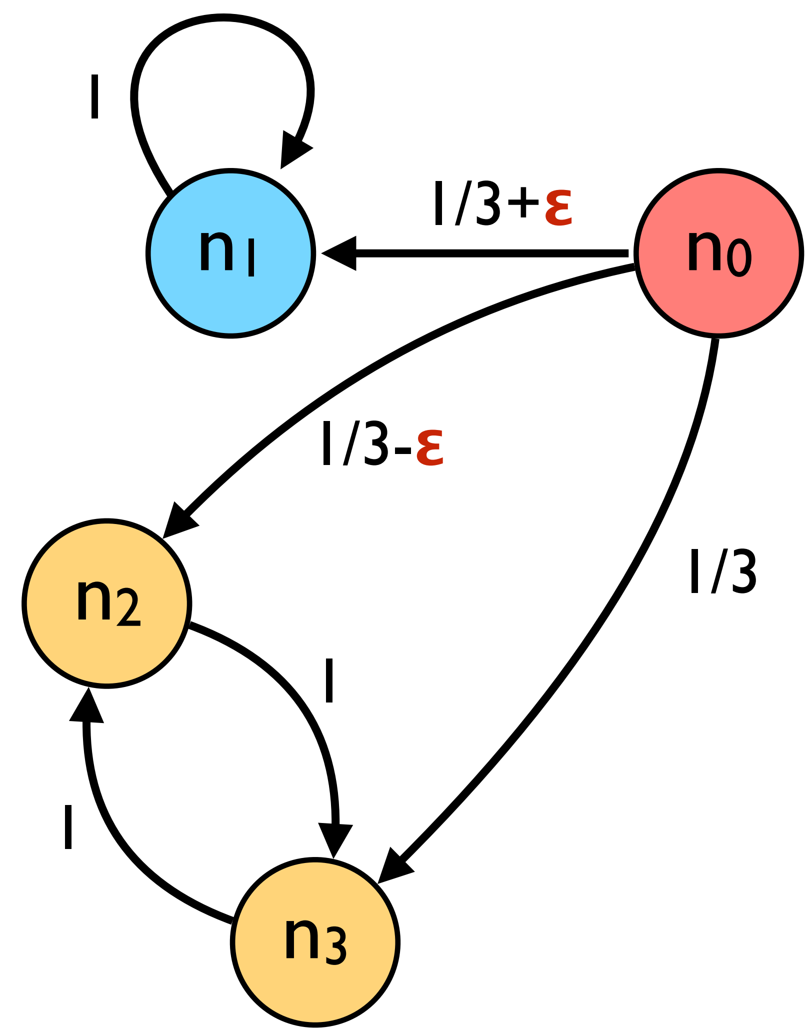


…but small variations may prevent aggregation

Bisimilarity Distance

$\mathcal{M} = (M, \tau, \ell, m_0)$ $\mathcal{N} = (N, \theta, \alpha, n_0)$

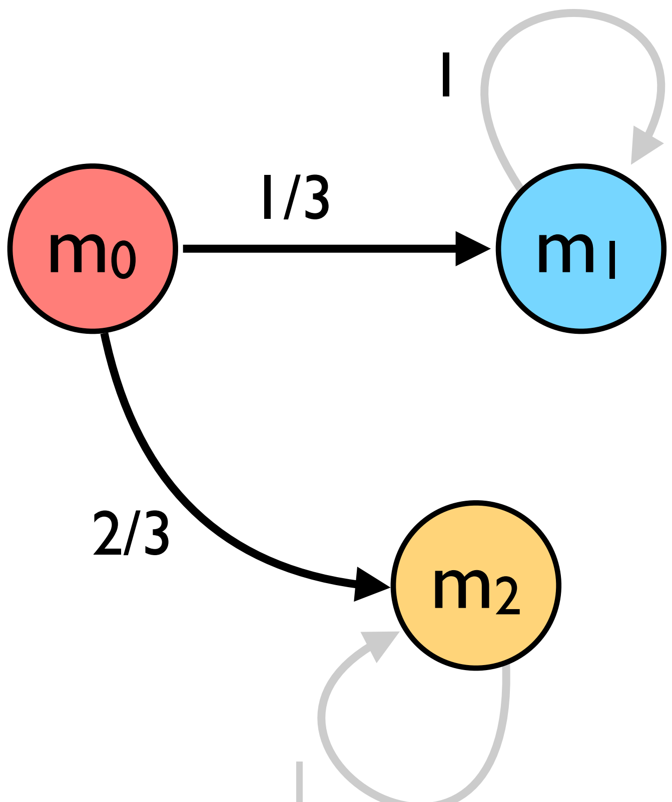


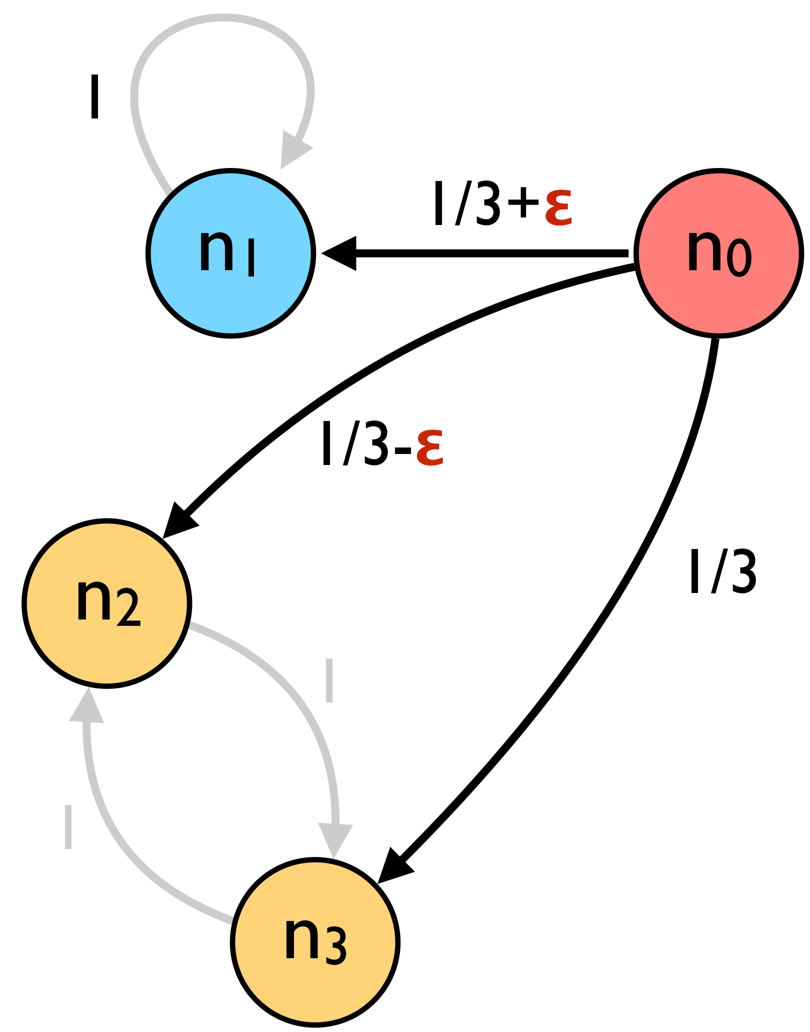


Bisimilarity Distance

$\mathcal{M} = (M, \tau, \ell, m_0)$

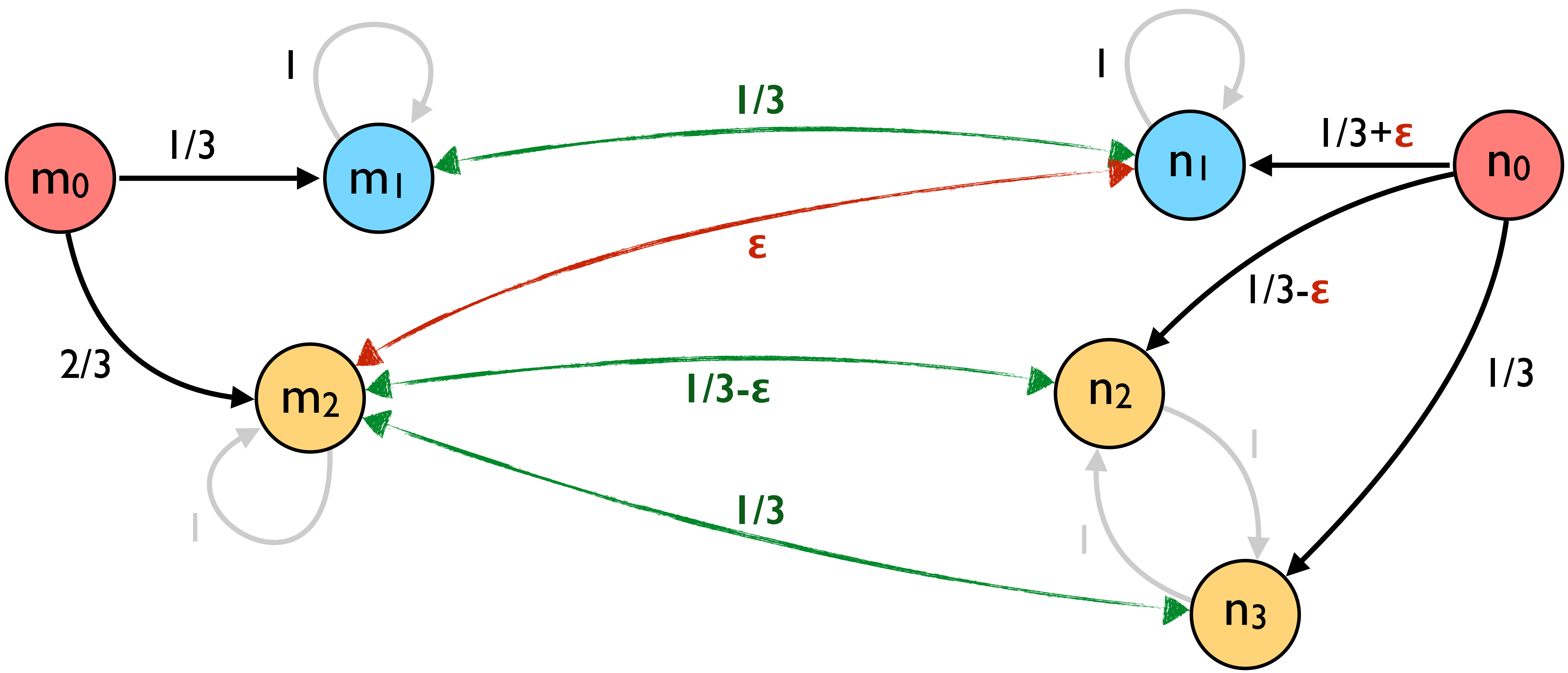
$\mathcal{N} = (N, \theta, \alpha, n_0)$





Bisimilarity Distance

$\mathcal{M} = (M, \tau, \ell, m_0)$ $\mathcal{N} = (N, \theta, \alpha, n_0)$



Bisimilarity Distance

(fixed point characterization by van Breugel & Worrell)

Given a parameter $\lambda \in (0,1]$, called *discount factor*, the *bisimilarity distance* δ_λ is the least fixed point of

$$\Delta_\lambda(d)(m,n) = \begin{cases} 1 & \text{if } \ell(m) \neq \alpha(n) \\ \lambda \cdot \mathcal{K}(d)(\tau(m),\theta(n)) & \text{otherwise} \end{cases}$$

discount at each step

Kantorovich lifting

coupling

$$\mathcal{K}(d)(\tau(m),\theta(n)) = \min \left\{ \sum d(u,v) \cdot C(u,v) \;\middle|\; \begin{array}{l} \sum_{u \in M} C(u,v) = \theta(n)(v) \\ \sum_{v \in N} C(u,v) = \tau(m)(u) \end{array} \right\}$$

Approximate verification

[Chen, van Breugel, Worrell - FoSSaCS'12]

$$|P(\mathcal{M})([\varphi]) - P(\mathcal{N})([\varphi])| \leq \delta_1(\mathcal{M},\mathcal{N})$$

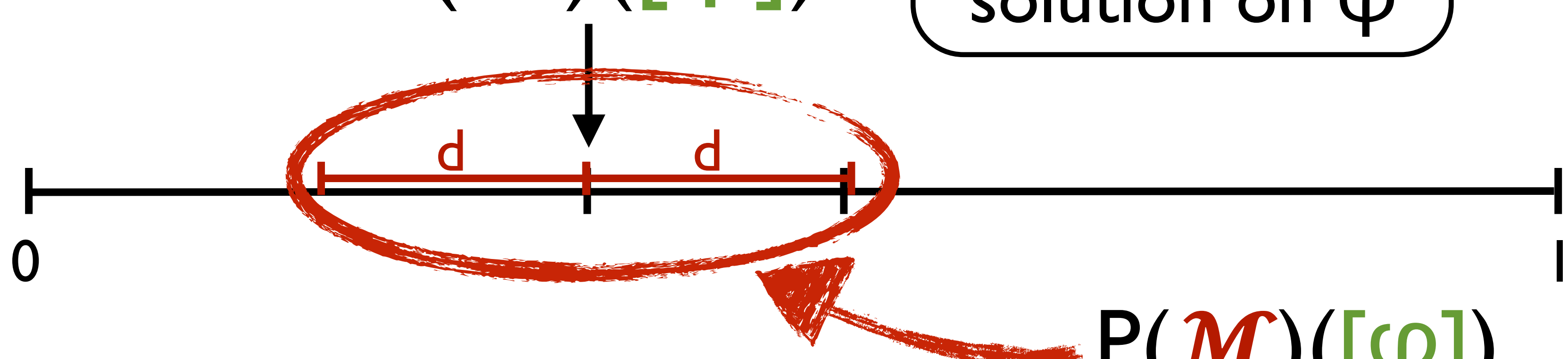
difference in the probability of satisfying φ

for all LTL formulas!

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

$P(\mathcal{N})([\varphi])$

approximate solution on φ



$P(\mathcal{M})([\varphi])$

Some natural questions

- Given an MC $\mathcal{M}$, and $k \in \mathbb{N}$, what is its closest k-state approximant?
- Does this always exist?
- Can we find one? How hard is it to get?

The CBA-λ problem

The Closest Bounded Approximant w.r.t. δ_λ

INSTANCE: An MC $\mathcal{M}$, and a positive integer k
OUTPUT: An MC $\mathcal{N}^*$ with at most k states minimizing $\delta_\lambda(\mathcal{M},\mathcal{N}^*)$

MCs with $\leq$ k states

$$\delta_\lambda(\mathcal{M},\mathcal{N}^*) = \inf \{ \delta_\lambda(\mathcal{M},\mathcal{N}) \mid \mathcal{N} \in MC(k)\}$$

generalization of bisimilarity quotient

CBA-λ has always a solution

$$\inf \{ \delta_\lambda(\mathcal{M},\mathcal{N}) \mid \mathcal{N} \in MC(k)\} =$$
$$= \inf\{ d(m_0,n_0) \mid \Delta_\lambda(d) \sqsubseteq d, \mathcal{N} \in MC(k) \}$$

Lemma (Meaningful labels)

For any $\mathcal{N}$'$\in$MC(k) there exists $\mathcal{N}\in$MC(k) with labels taken from $\mathcal{M}$, such that $\delta_\lambda(\mathcal{M},\mathcal{N}) \leq \delta_\lambda(\mathcal{M},\mathcal{N}')$.

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 bilinear program

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

The complexity of CBA-λ

We study its complexity by looking at its decision variant

The Bounded Approximant threshold w.r.t. δ_λ

INSTANCE: An MC $\mathcal{M}$, a positive integer k, and a rational bound ε

OUTPUT: yes iff $\delta_\lambda(\mathcal{M},\mathcal{N}) \leq \varepsilon$ for some $\mathcal{N} \in MC(k)$

Theorem:

For any $\lambda \in (0,1]$, BA-λ is in PSPACE

proof sketch: *we can encode the question $\langle \mathcal{M}, k, \varepsilon \rangle \in$ BA-λ to that of asking for the feasibility of a set of bilinear inequalities. This is a decision problem in for the existential theory of the reals, thus it can be solved in PSPACE* [Canny - STOC'88].

The complexity of CBA-λ

We study its complexity by looking at its decision variant

The Bounded Approximant threshold w.r.t. δ_λ

INSTANCE: An MC $\mathcal{M}$, a positive integer k, and a rational bound ε

OUTPUT: yes iff $\delta_\lambda(\mathcal{M},\mathcal{N}) \leq \varepsilon$ for some $\mathcal{N} \in \mathrm{MC}(k)$

Theorem:

For any $\lambda \in (0,1]$, BA-λ is in PSPACE

Theorem:

For any $\lambda \in (0,1]$, BA-λ is NP-hard

proof idea: *by reduction from VERTEX COVER*

…the hardness of BA-λ opens a new question:
is it easy to choose a "good" bound k?

$\mathcal{N}$ is a **significant approximant** if
$\delta_\lambda(\mathcal{M},\mathcal{N}) < 1$

The MSAB-λ problem

The Minimum Significant Approximant Bound w.r.t. δ_λ

INSTANCE: An MC $\mathcal{M}$, and a positive integer k

OUTPUT: The smallest k such that $\delta_\lambda(\mathcal{M},\mathcal{N})<1$, for some $\mathcal{N} \in \mathrm{MC}(k)$

The Significant Bounded Appr

INSTANCE: An MC $\mathcal{M}$, and a p

OUTPUT: yes iff $\delta_\lambda(\mathcal{M},\mathcal{N})<1$ for some $\mathcal{N}$

NP-complete for $\lambda=1$

A practical solution: **EM** Algorithm

- Given $\mathcal{M}$ and a significant approximant $\mathcal{N}_0$
- it produces a sequence $\mathcal{N}_0, \ldots, \mathcal{N}_h$ having successively decreased distance from $\mathcal{M}$
- $\mathcal{N}_h$ is a sub-optimal solution of CBA-λ

Intuitive idea:

assign greater probability to transitions that are most representative of the behavior of $\mathcal{M}$

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.

What we have seen

Theoretical Results

We studied **metric-based state space reduction for MCs**

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

Practical Results

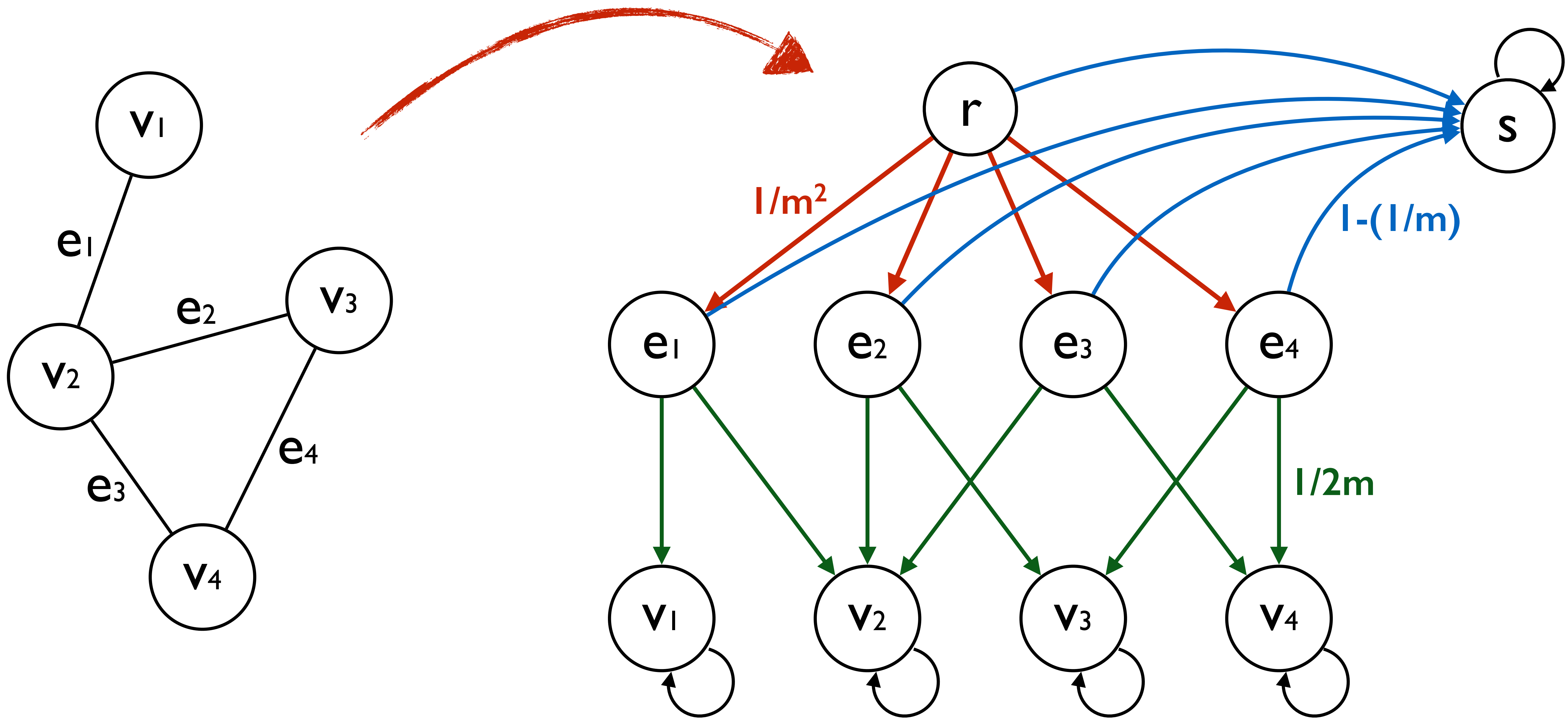
We proposed an EM method to obtain a sub-optimal approximants

Ongoing & Future work

- Improve the encoding as bilinear program
- Study the CBA problem w.r.t. other
  - behavioral distances (e.g. Total Variation)
  - models (e.g. MDP, CTMC, Prob. Automata)

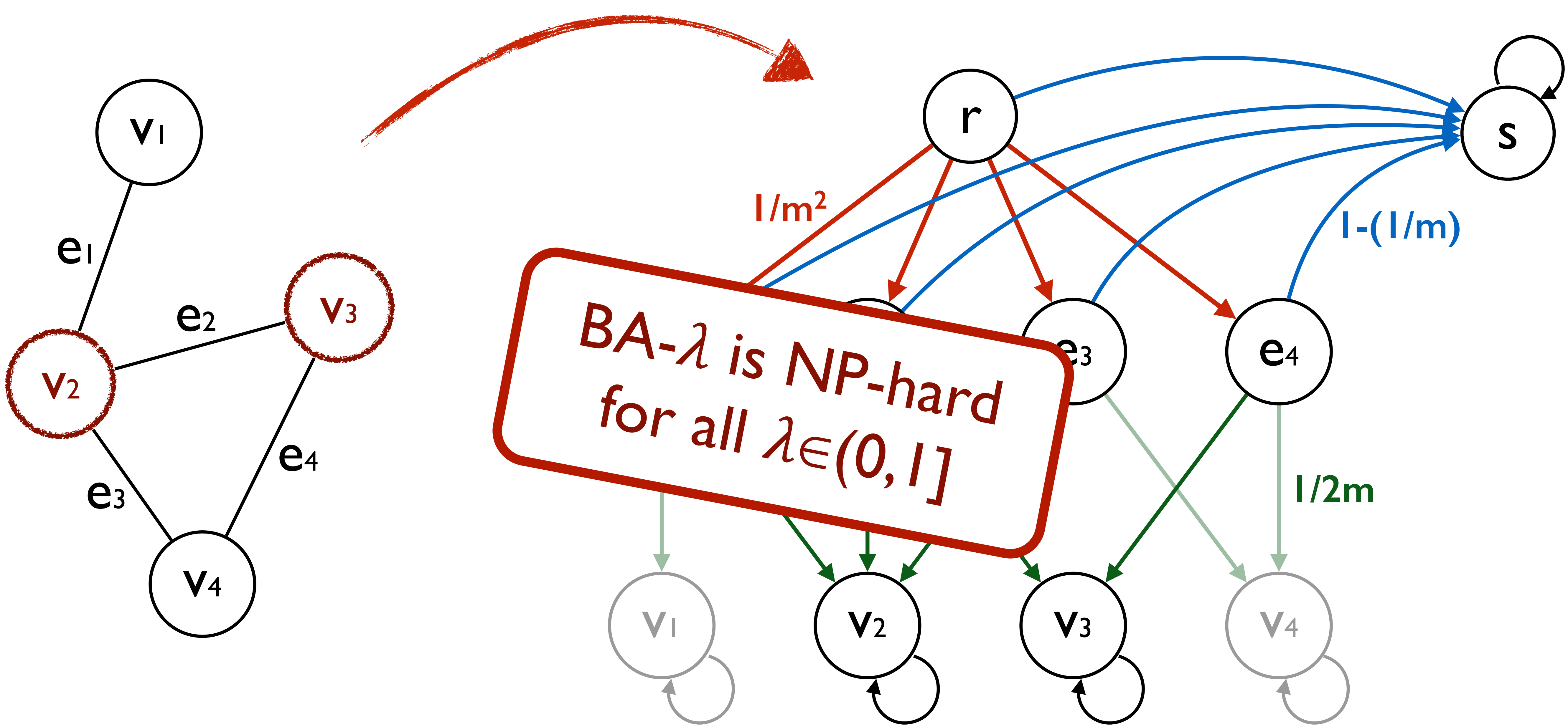
Appendix

Vertex Cover $\leq_P$ BA-λ



$$\langle G,h\rangle \in \text{VertexCover} \iff \langle \mathcal{M}_G,\ m+h+2,\ \lambda^2/2m^2\rangle \in \text{BA-}\lambda$$

Vertex Cover $\leq_P$ BA-λ



$$\langle G, h \rangle \in \text{VertexCover} \iff \langle \mathcal{M}_G,\ m+h+2,\ \lambda^2/2m^2 \rangle \in \text{BA-}\lambda$$