On the Total Variation Distance of SMCs

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

17 December 2014 - Bologna, Italy

FOCUS seminars

Before to start...

Given $\mu,\nu : \Sigma \to \mathbb{R}_+$ measures on (X,Σ)

Total Variation Distance

$$||\mu - \nu|| = \sup_{E \in \Sigma} |\mu(E) - \nu(E)|$$

Before to start...

Given $\mu, \nu : \Sigma \to \mathbb{R}_+$ measures on (X, Σ)

Total Variation Distance

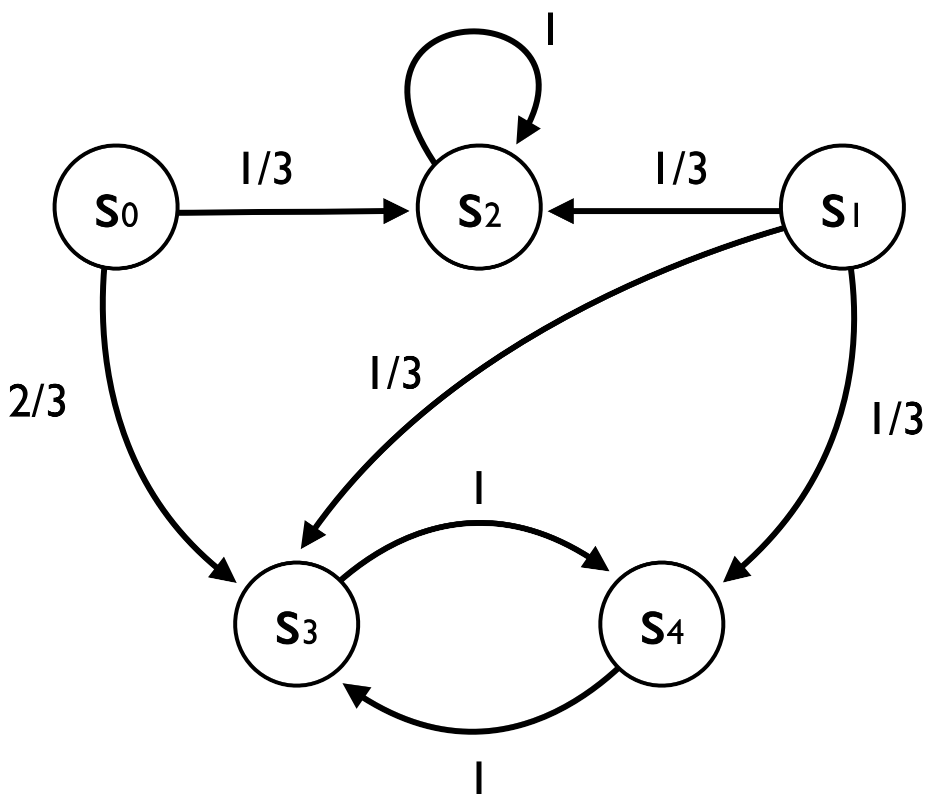
$$||\mu - \nu|| = \sup_{E \in \Sigma} |\mu(E) - \nu(E)|$$

The largest possible difference that μ and ν assign to the same event

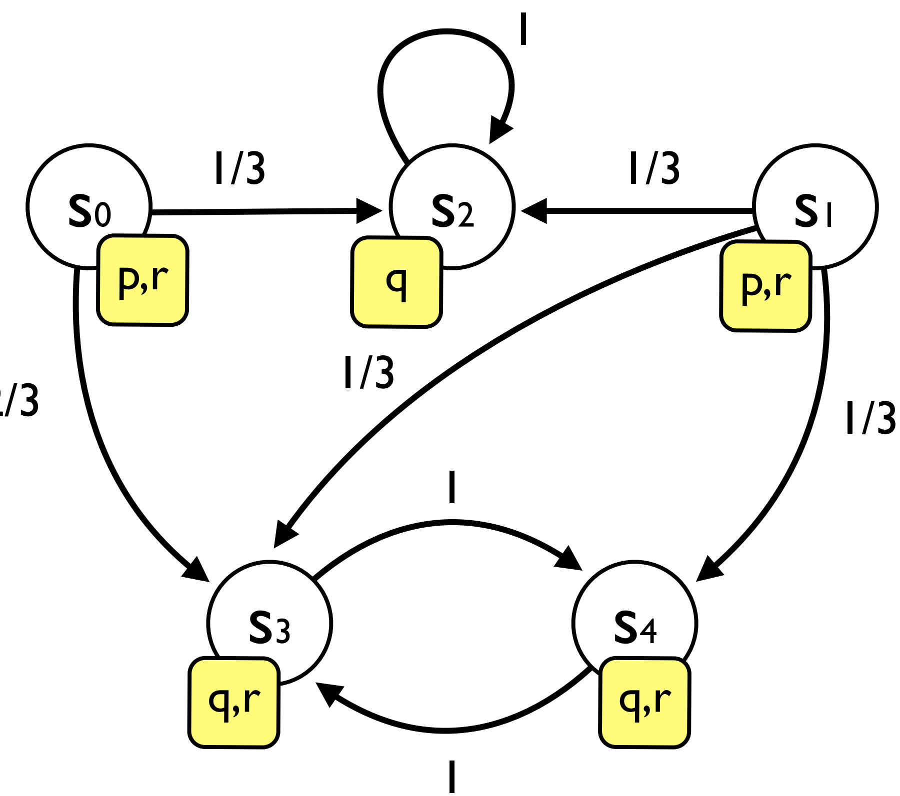
Outline

- Semi-Markov Chains (SMCs)
- Total Variation vs Model Checking of SMCs
- An Approximation Algorithm
- Concluding Remarks

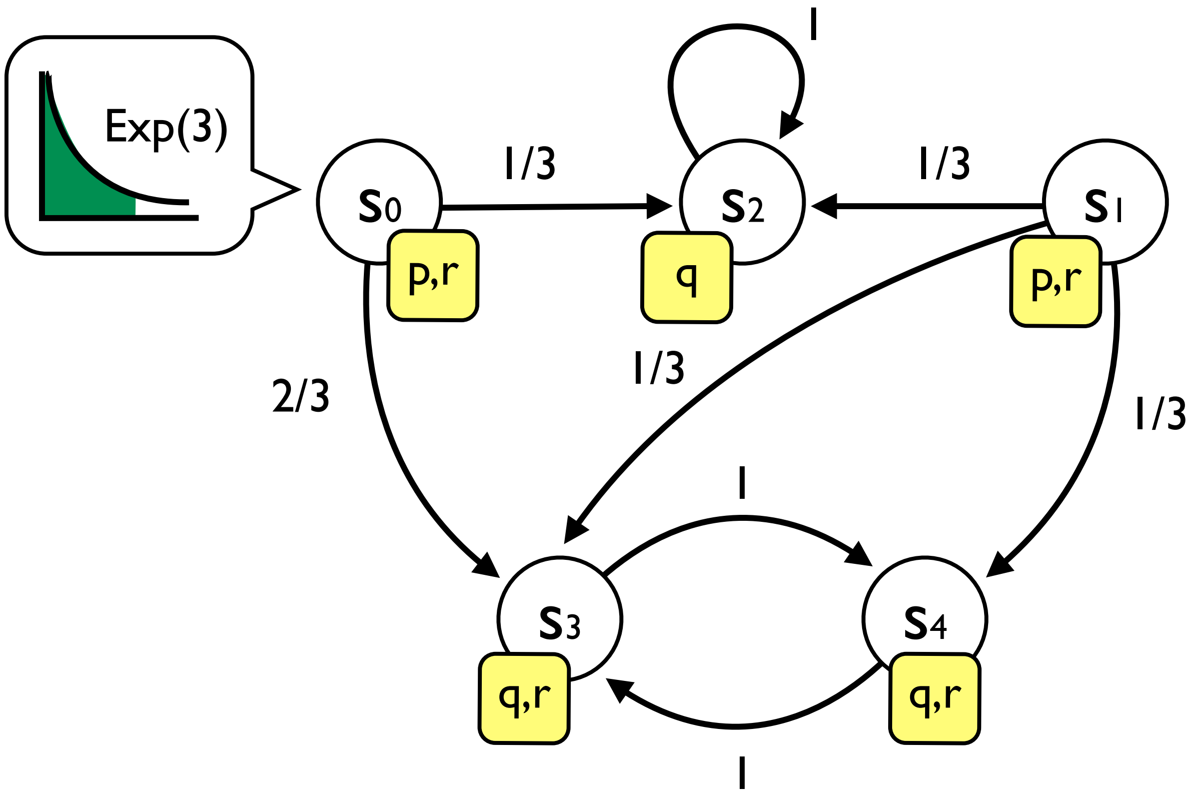
semi-Markov Chains



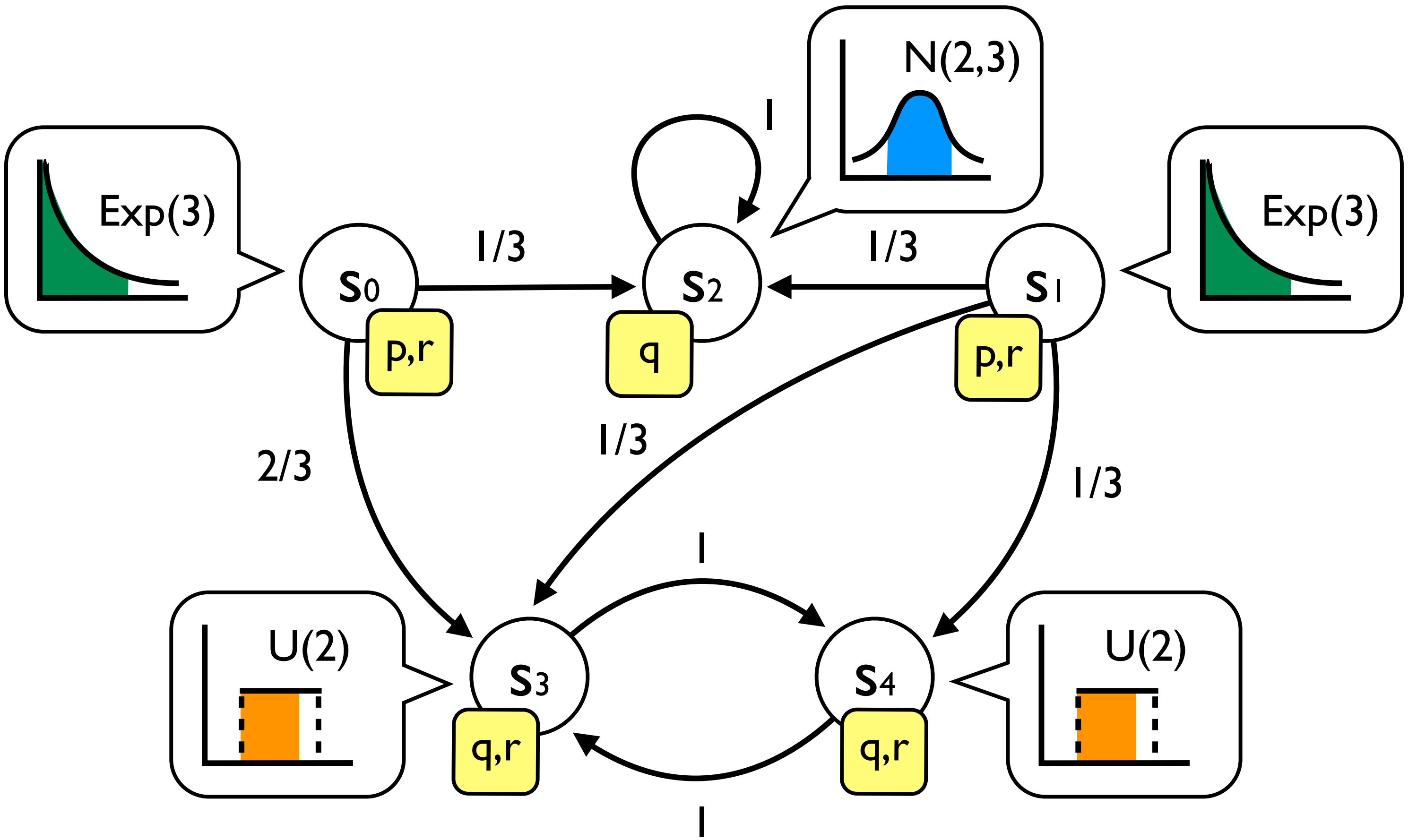
semi-Markov Chains



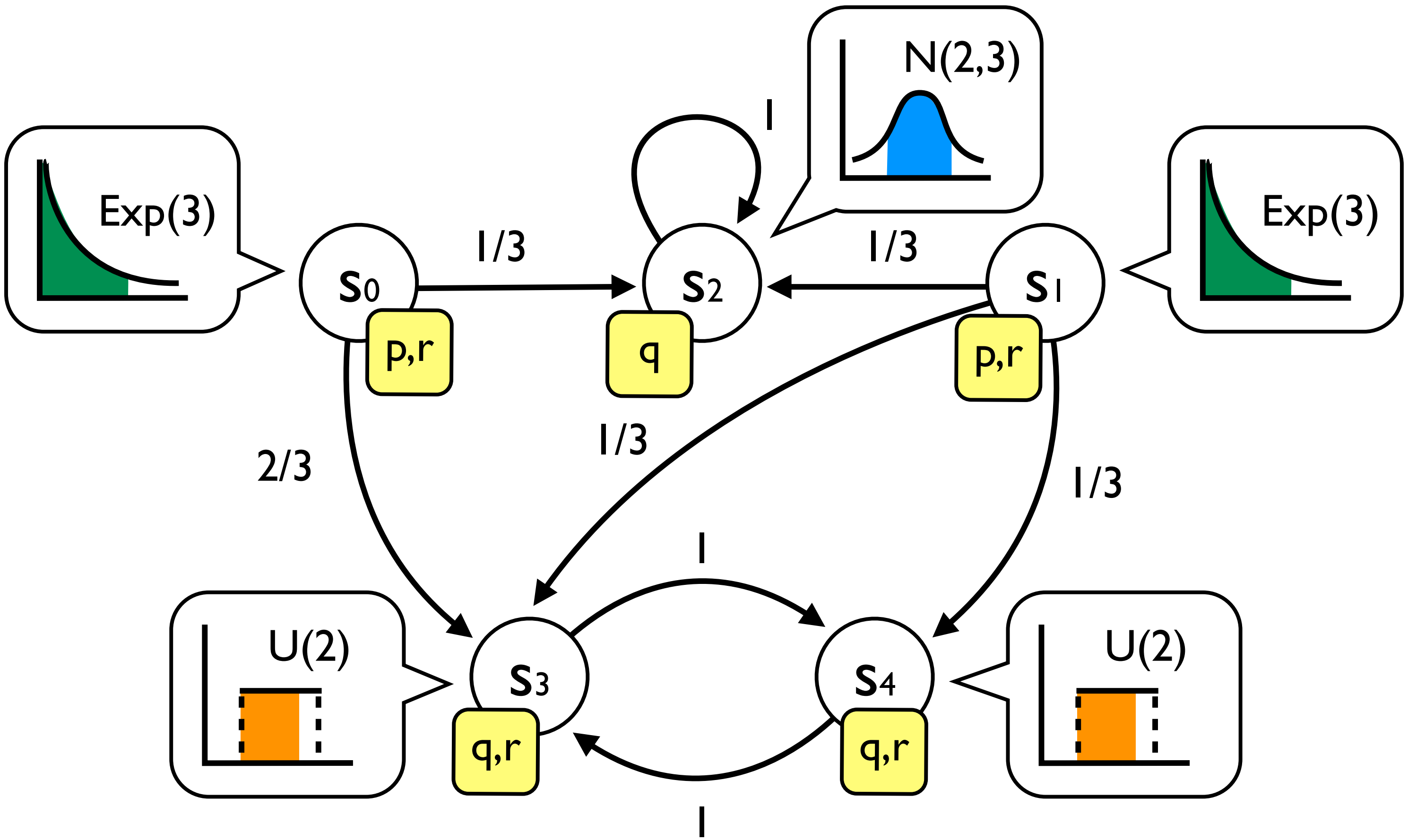
semi-Markov Chains



semi-Markov Chains

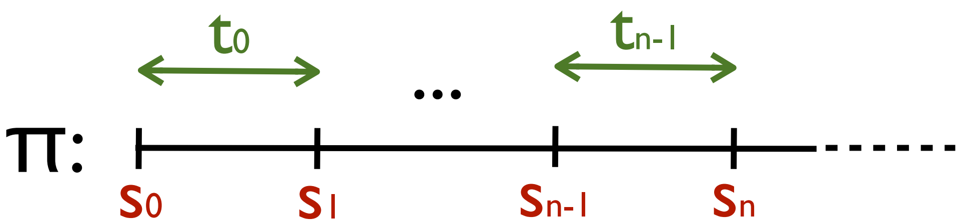


semi-Markov Chains



Given an initial state, SMCs can be interpreted as "machines" that emit timed traces of states with a certain probability

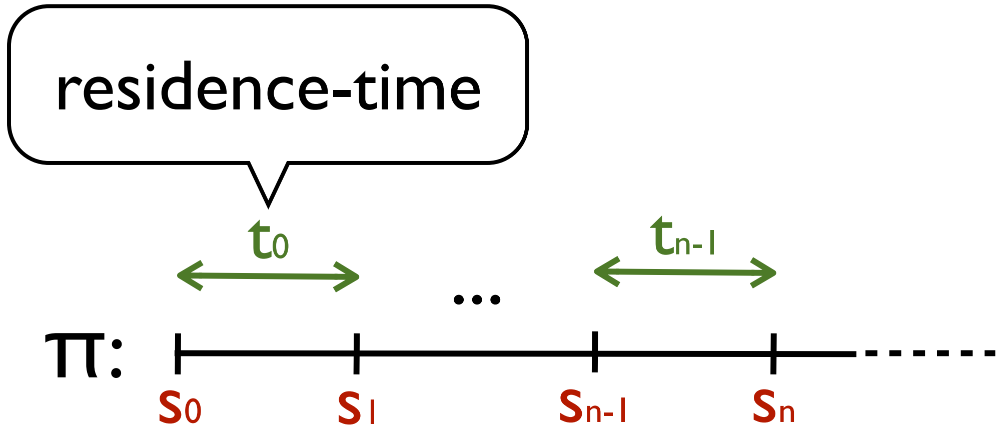
Timed paths & Events



$\in \mathfrak{C}(S_0, R_0, \ldots, R_{n-1}, S_n)$

$P[s](\mathfrak{C}(S_0, R_0, \ldots, R_{n-1}, S_n)) =$ "probability that, *starting from s*, the SMC emits a timed path with prefix in $S_0 \times R_0 \times \ldots \times R_{n-1} \times S_n$"

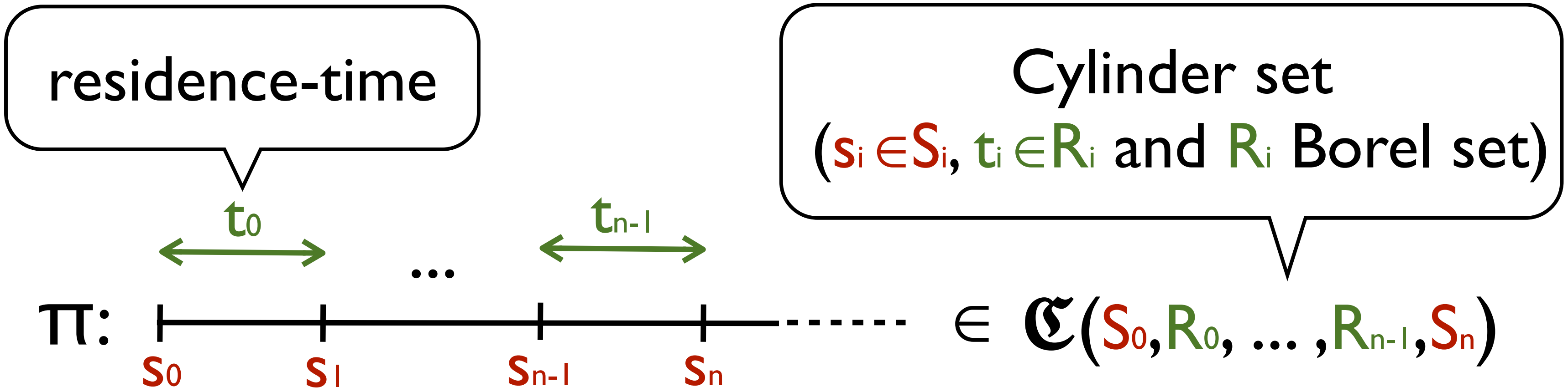
Timed paths & Events



$\in \mathfrak{C}(S_0, R_0, \ldots, R_{n-1}, S_n)$

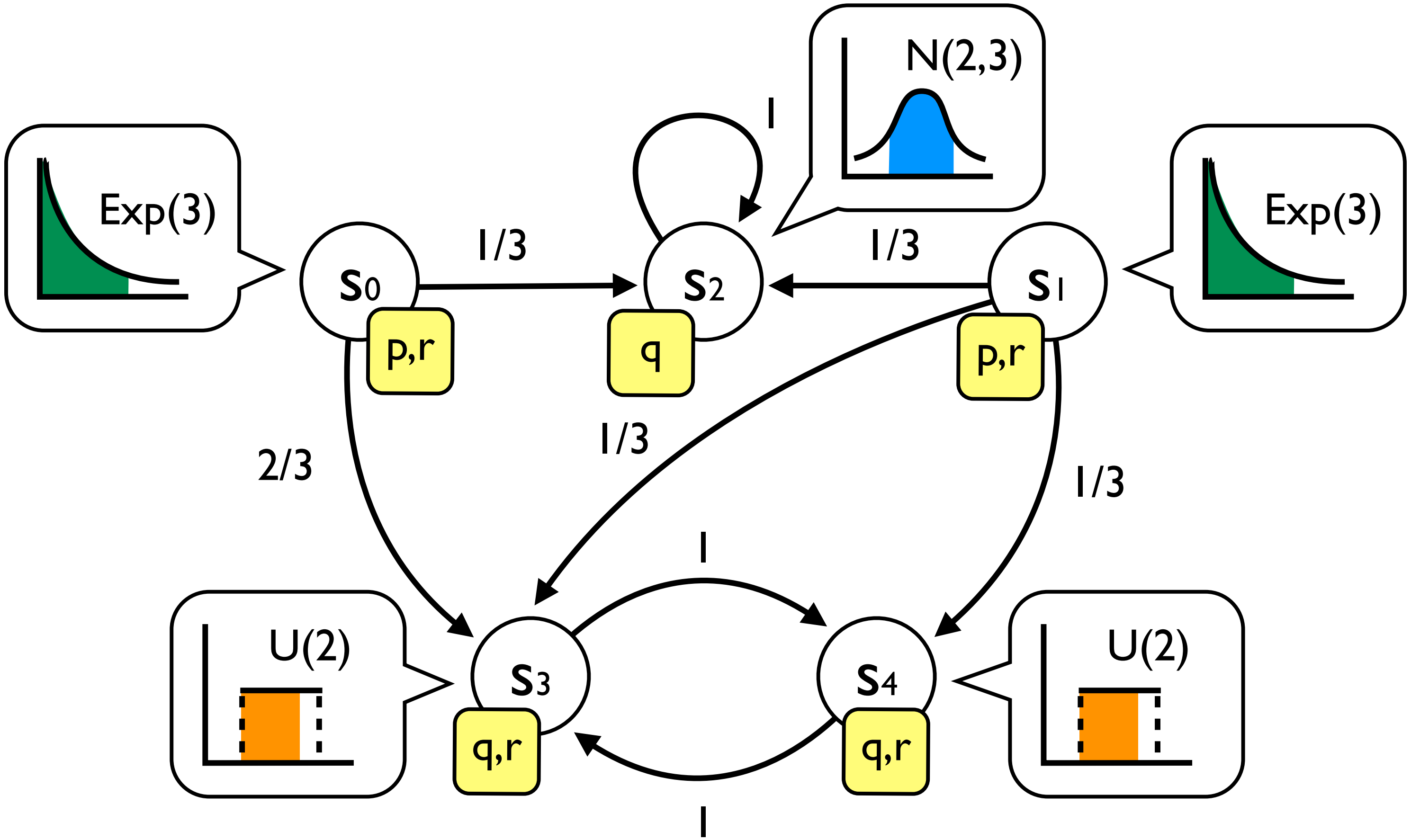
$P[s](\mathfrak{C}(S_0, R_0, \ldots, R_{n-1}, S_n)) =$ "probability that, *starting from s*, the SMC emits a timed path with prefix in $S_0 \times R_0 \times \ldots \times R_{n-1} \times S_n$"

Timed paths & Events

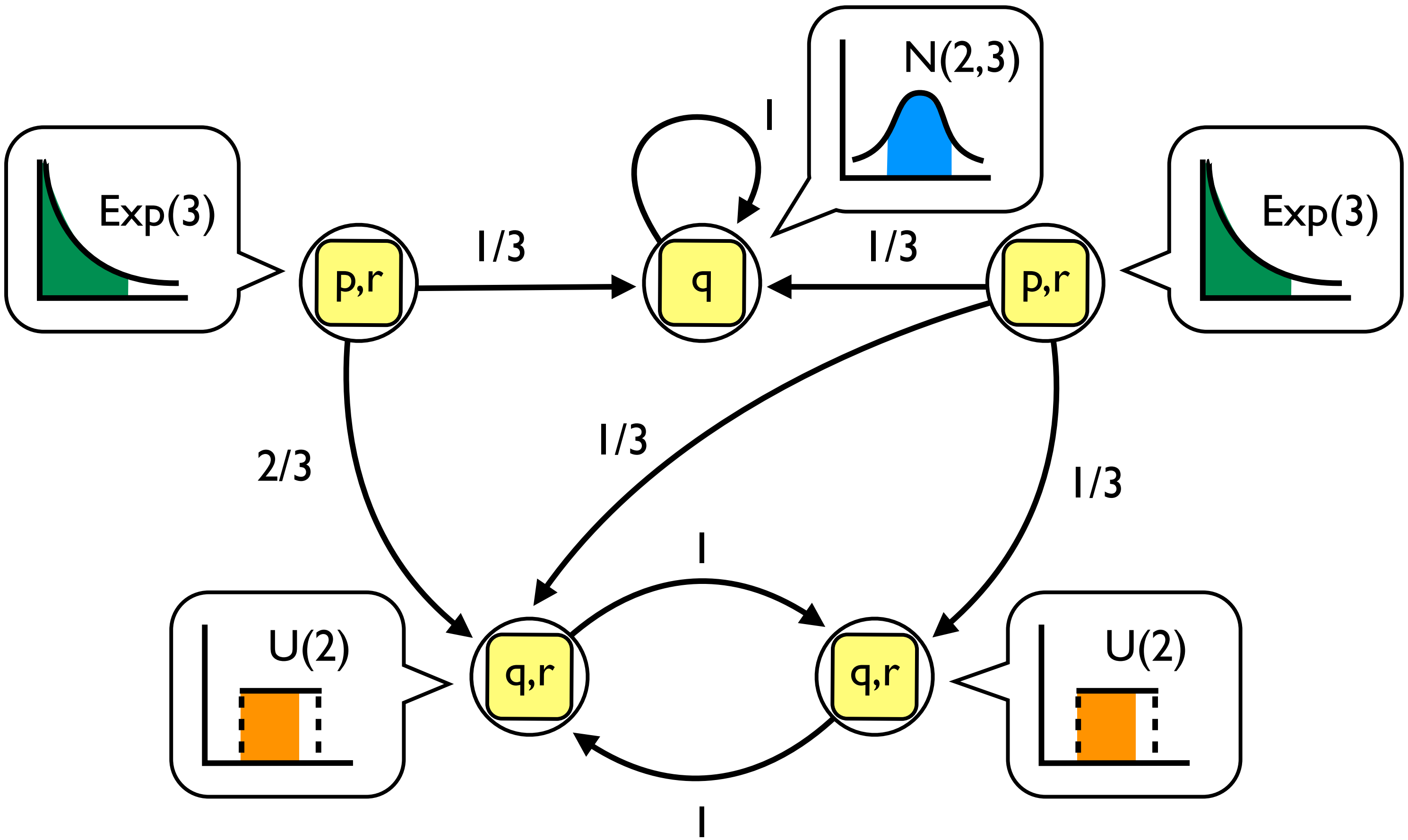


$P[s](\mathfrak{C}(S_0, R_0, \ldots, R_{n-1}, S_n)) =$ "probability that, *starting from s*, the SMC emits a timed path with prefix in $S_0 \times R_0 \times \ldots \times R_{n-1} \times S_n$"

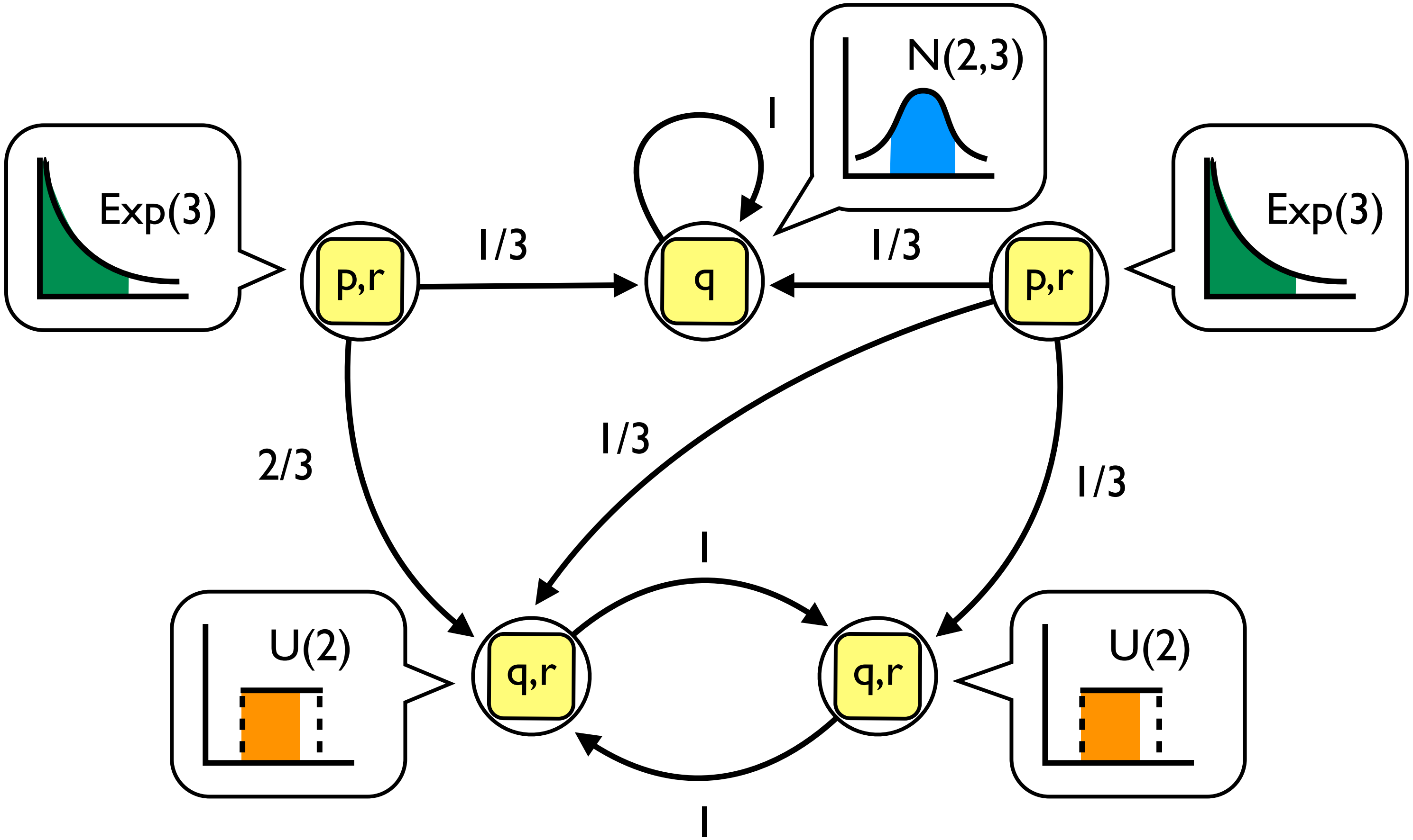
Probabilistic Trace Equiv.



Probabilistic Trace Equiv.

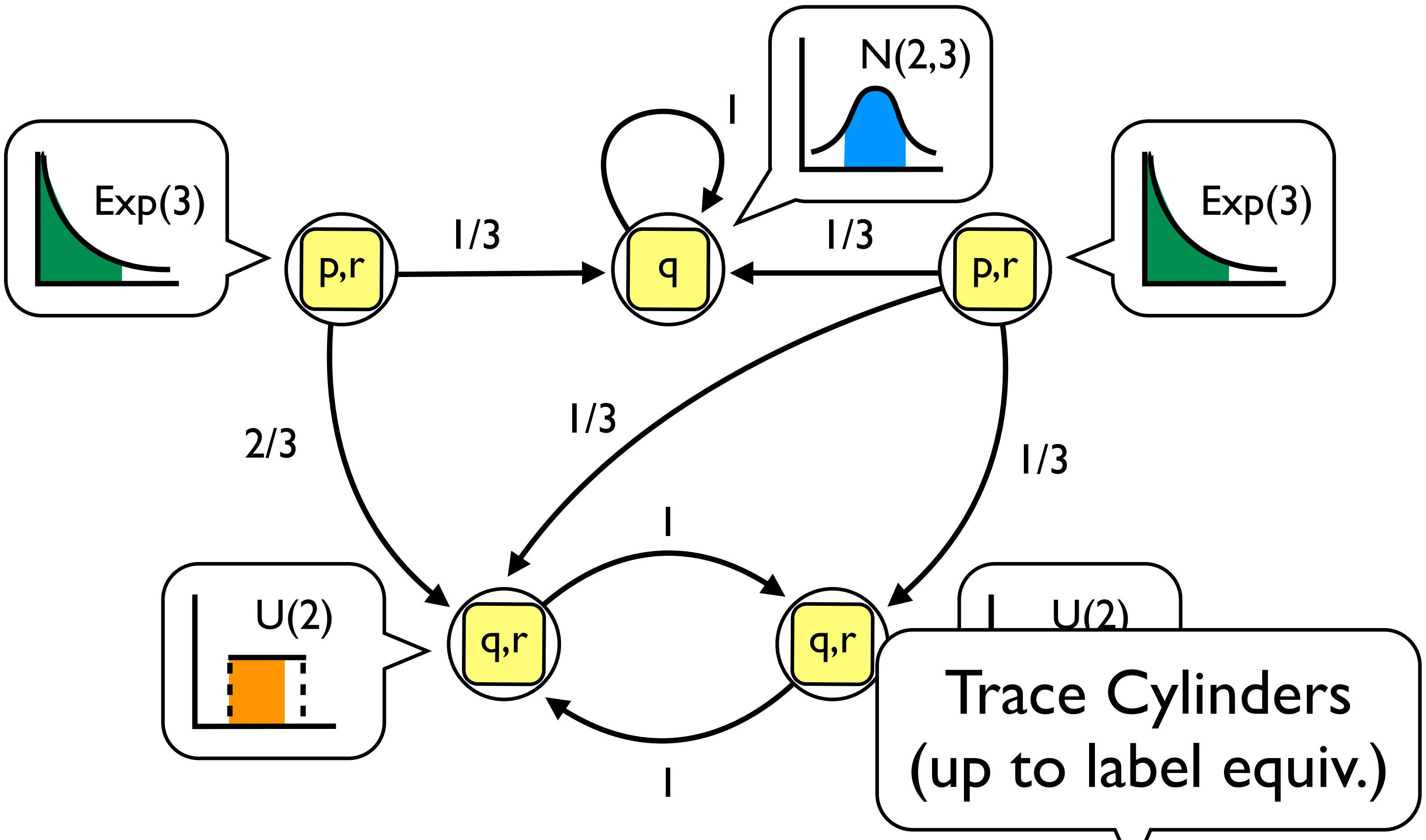


Probabilistic Trace Equiv.



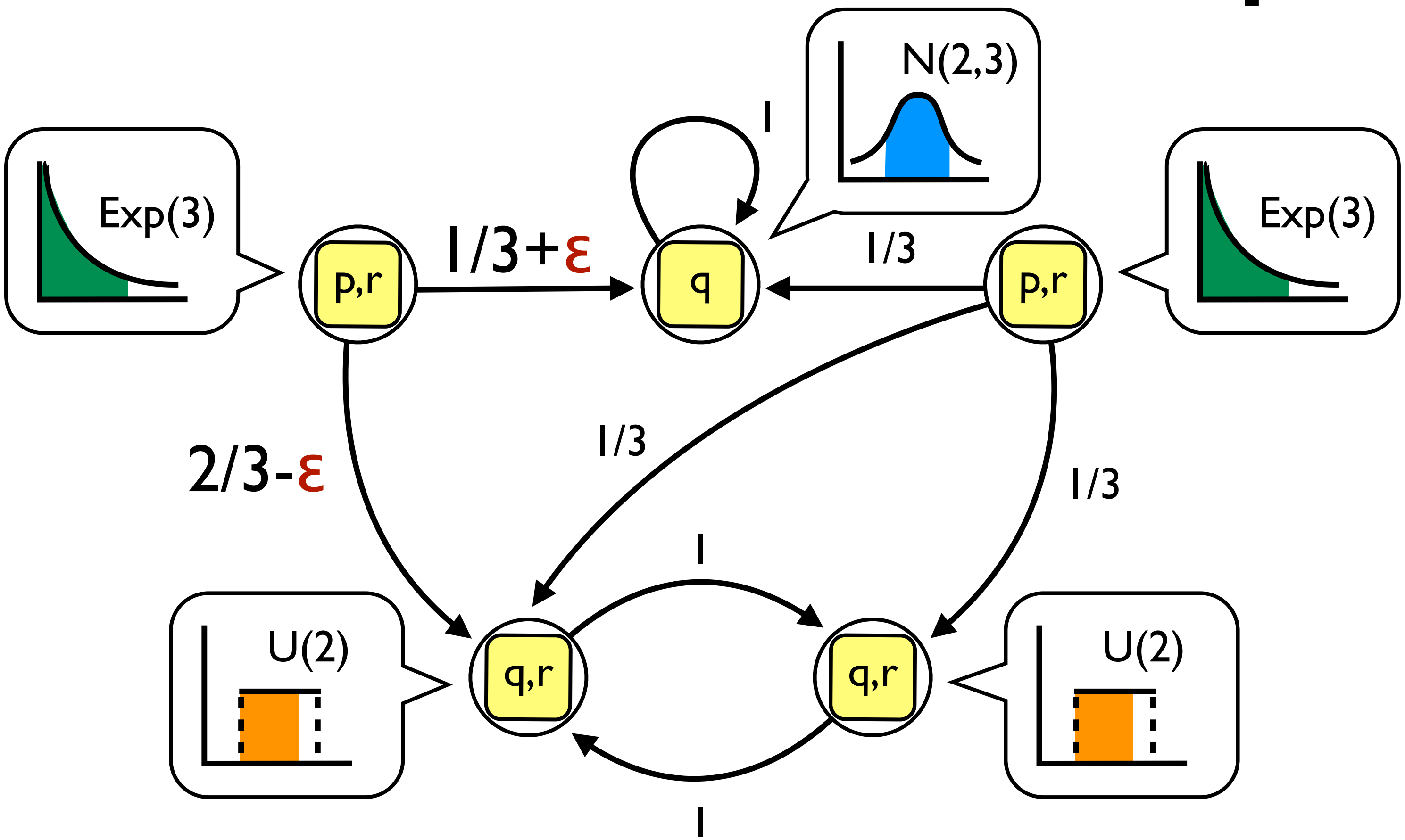
$$P[s_0](\mathfrak{C}(L_0, R_0, \ldots, R_{n-1}, L_n)) = P[s_1](\mathfrak{C}(L_0, R_0, \ldots, R_{n-1}, L_n))$$

Probabilistic Trace Equiv.



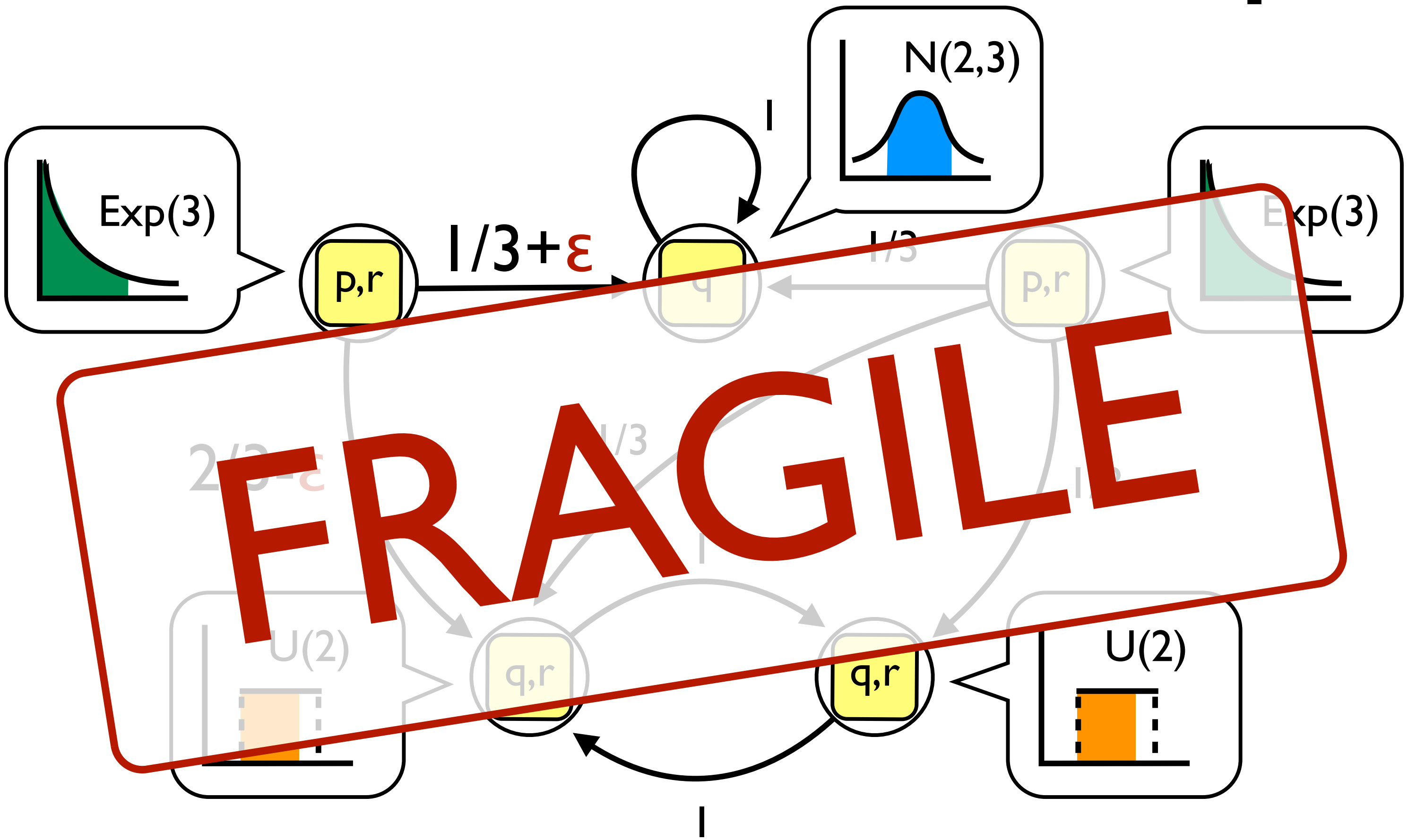
$$P[s_0](\mathfrak{C}(L_0, R_0, \ldots, R_{n-1}, L_n)) = P[s_1](\mathfrak{C}(L_0, R_0, \ldots, R_{n-1}, L_n))$$

Probabilistic Trace Equiv.



$$P[s_0](\mathfrak{C}(\boxed{p,r}, \mathbb{R}, \boxed{q},)) = 1/3+\varepsilon \neq 1/3 = P[s_1](\mathfrak{C}(\boxed{p,r}, \mathbb{R}, \boxed{q},))$$

Probabilistic Trace Equiv.



$P[s_0](\mathfrak{C}(\boxed{p,r}, \mathbb{R}, \boxed{q},)) = 1/3+\varepsilon \neq 1/3 = P[s_1](\mathfrak{C}(\boxed{p,r}, \mathbb{R}, \boxed{q},))$

Trace Pseudometric

$$d(s,s') = \sup_{E \in \sigma(\mathcal{T})} |P[s](E) - P[s'](E)|$$

σ-algebra generated from Trace Cylinders

Trace Pseudometric

$$d(s,s') = \sup_{E \in \sigma(\mathcal{T})} |P[s](E) - P[s'](E)|$$

σ-algebra generated from Trace Cylinders

It's a Behavioral Distance!

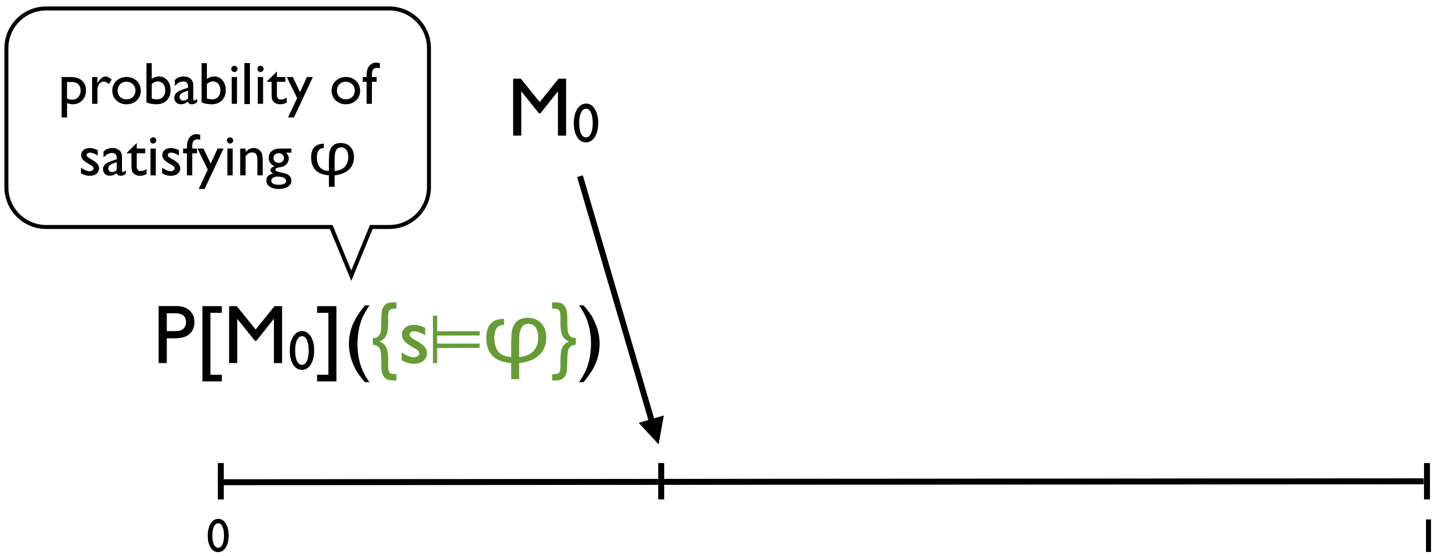
$$d(s,s') = 0 \quad \text{iff} \quad s \approx_T s'$$

Distance = Approx. Error

Distance $\stackrel{?}{=}$ Approx. Error

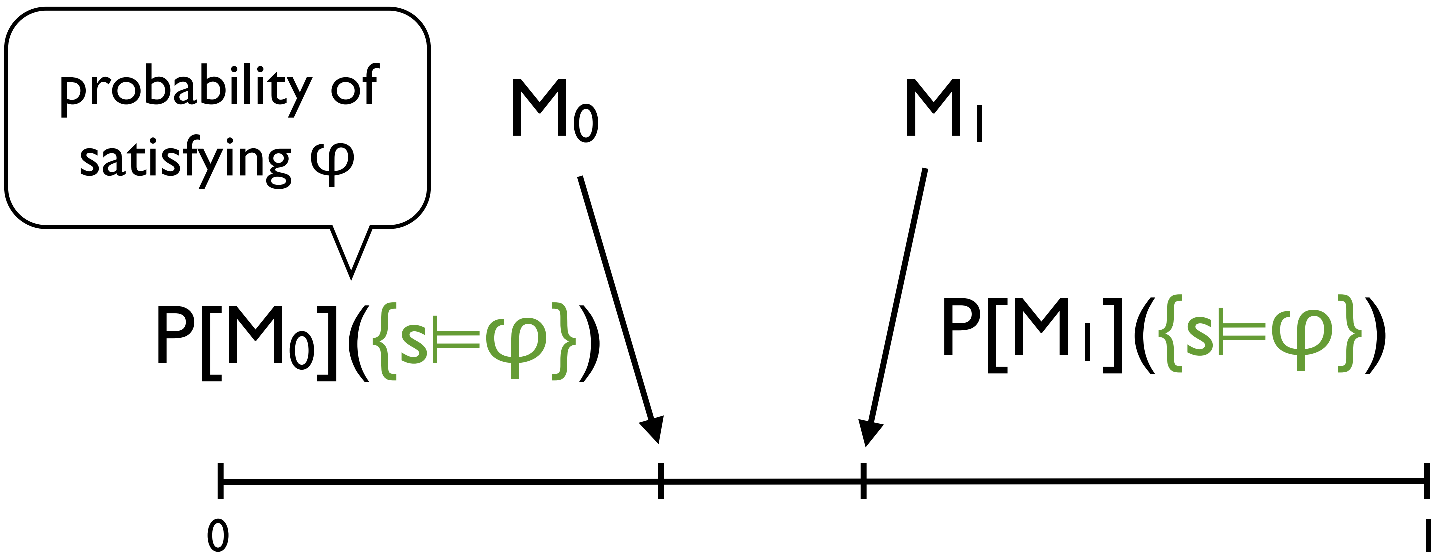
Distance $\stackrel{?}{=}$ Approx. Error

Example: Probabilistic Model Checking



Distance $\stackrel{?}{=}$ Approx. Error

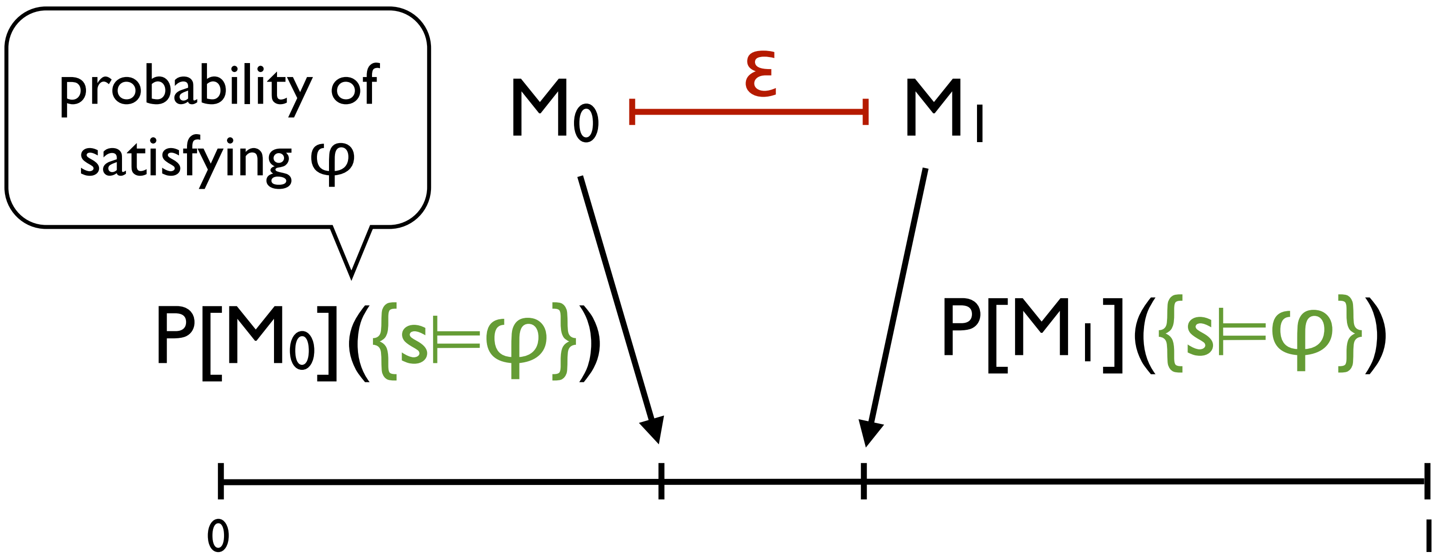
Example: Probabilistic Model Checking



$$|P[M_0](\{s \models \varphi\}) - P[M_0](\{s \models \varphi\})|$$

Distance $\stackrel{?}{=}$ Approx. Error

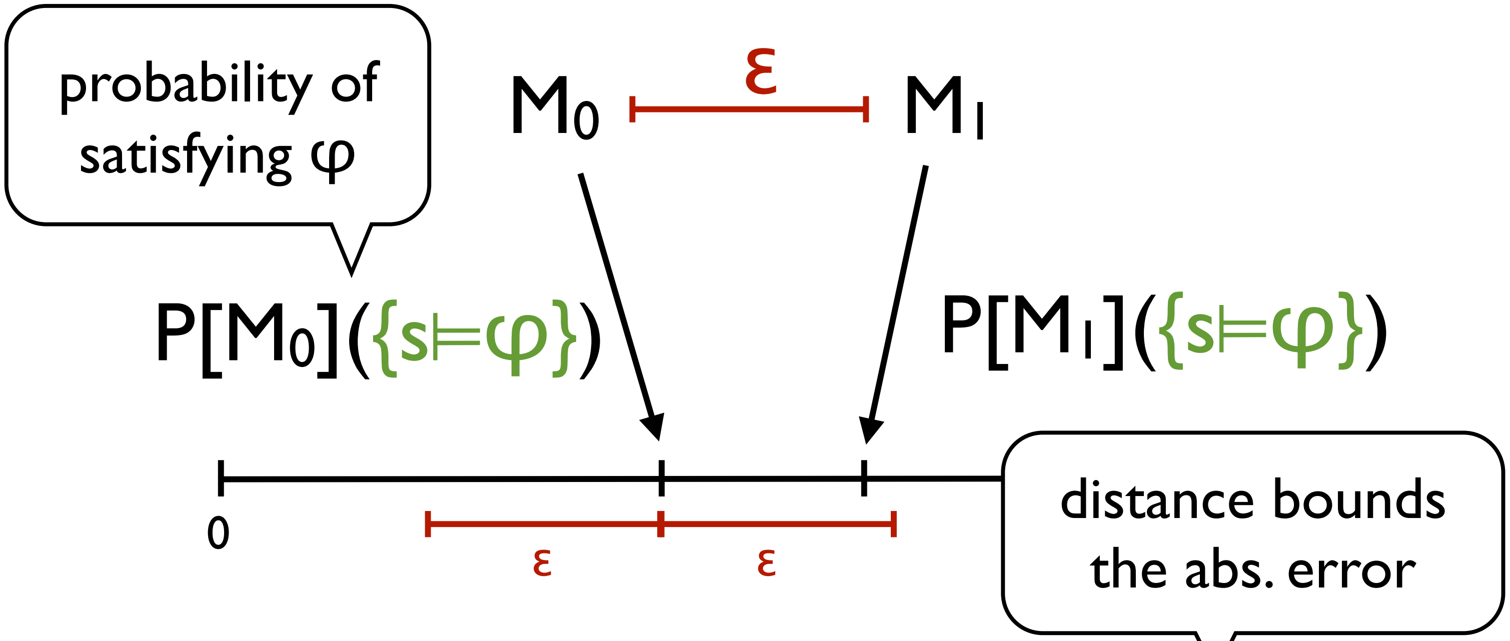
Example: Probabilistic Model Checking



$$|P[M_0](\{s \models \varphi\}) - P[M_0](\{s \models \varphi\})|$$

Distance $\stackrel{?}{=}$ Approx. Error

Example: Probabilistic Model Checking



$$|P[M_0](\{s \models \varphi\}) - P[M_0](\{s \models \varphi\})| \leq \varepsilon$$

Trace Distance
vs.
Model Checking

(i.e., does it provide a good approximation error?)

Model Checking SMCs

i.e., measuring the likelihood that a
a linear real-time property is satisfied by the SMC

SMC $\models$ Linear Real-time Spec.

Model Checking SMCs

i.e., measuring the likelihood that a
a linear real-time property is satisfied by the SMC

SMC $\models$ Linear Real-time Spec.

represented as
Metric Temporal Logic
formulas

Model Checking SMCs

i.e., measuring the likelihood that a
a linear real-time property is satisfied by the SMC

SMC $\models$ Linear Real-time Spec.

represented as
Metric Temporal Logic
formulas

... or languages
recognized
by Timed Automata

Model Checking SMCs

i.e., measuring the likelihood that a
a linear real-time property is satisfied by the SMC

a proper measurable set!

SMC $\models$ Linear Real-time Spec.

represented as
Metric Temporal Logic
formulas

... or languages
recognized
by Timed Automata

(Alur-Henzinger)

Metric Temporal Logic

Next — Until

$$\varphi \coloneqq p \mid \bot \mid \varphi \rightarrow \varphi \mid X^{I}\varphi \mid \varphi U^{I} \varphi$$

(*) $I \subseteq \mathbb{R}$ closed interval with *rational* endpoints

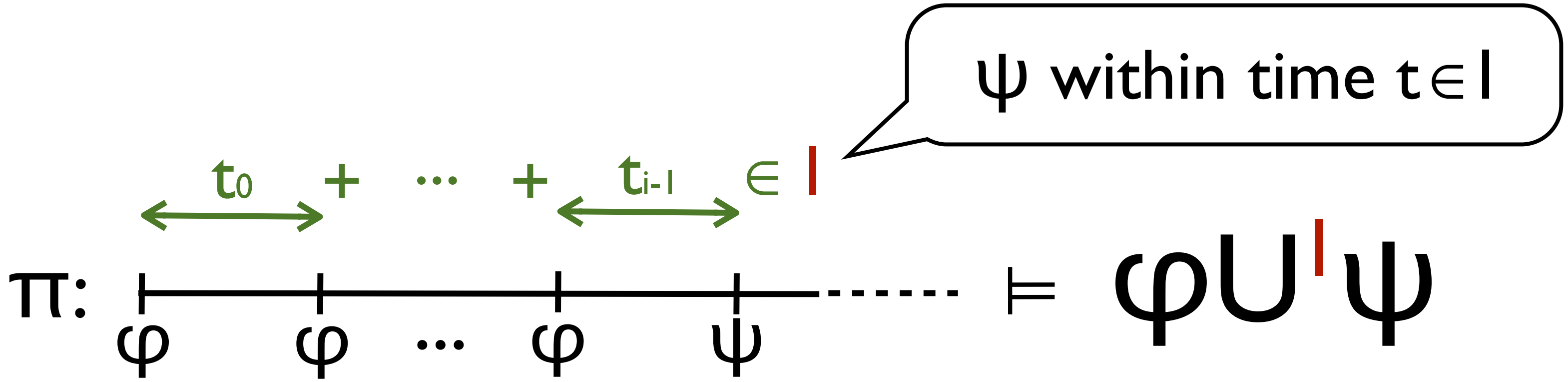
(Alur-Henzinger)

Metric Temporal Logic

Next / Until

$$\varphi ::= p \mid \bot \mid \varphi \rightarrow \varphi \mid X^{I}\varphi \mid \varphi U^{I} \varphi$$

(*) $I \subseteq \mathbb{R}$ closed interval with *rational* endpoints



MTL distance

(max error w.r.t. MTL properties)

set of timed paths that satisfy φ

$$MTL(s,s') = \sup_{\varphi \in MTL} |P[s](\{\pi \vDash \varphi\}) - P[s'](\{\pi \vDash \varphi\})|$$

MTL distance

(max error w.r.t. MTL properties)

set of timed paths that satisfy φ

measurable in $\sigma(\mathcal{T})$

$$MTL(s,s') = \sup_{\varphi \in MTL} |P[s](\{\pi \models \varphi\}) - P[s'](\{\pi \models \varphi\})|$$

Relation with Trace Distance

$$MTL(s,s') \leq d(s,s') = \sup_{E \in \sigma(\mathcal{T})} |P[s](E) - P[s'](E)|$$

MTL distance

(max error w.r.t. MTL properties)

set of timed paths that satisfy φ

measurable in $\sigma(\mathcal{T})$

$$MTL(s,s') = \sup_{\varphi \in MTL} |P[s](\{\pi \models \varphi\}) - P[s'](\{\pi \models \varphi\})|$$

Relation with Trace Distance

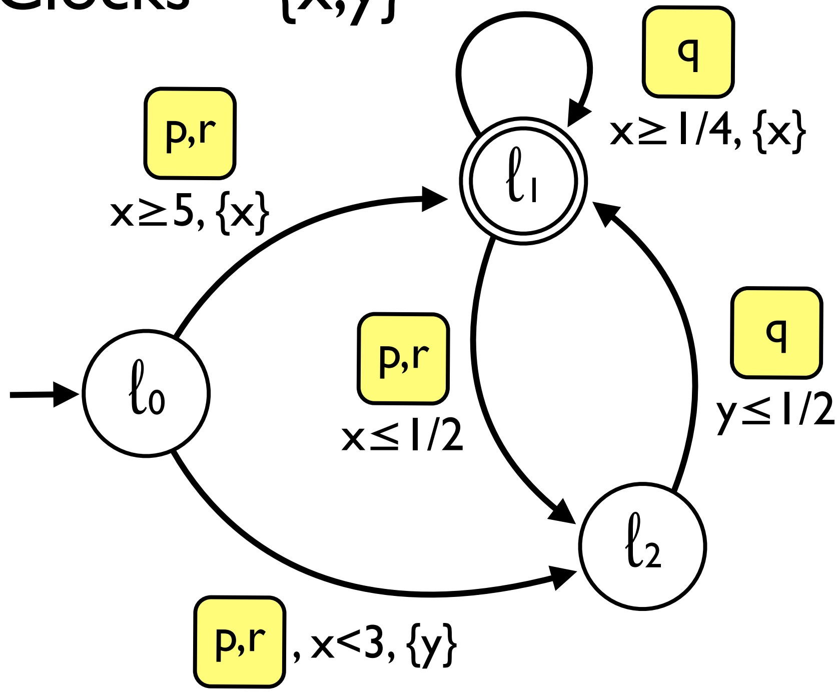
$$MTL(s,s') = d(s,s') = \sup_{E \in \sigma(\mathcal{T})} |P[s](E) - P[s'](E)|$$

(Alur-Dill)

(Muller) Timed Automata

without invariants

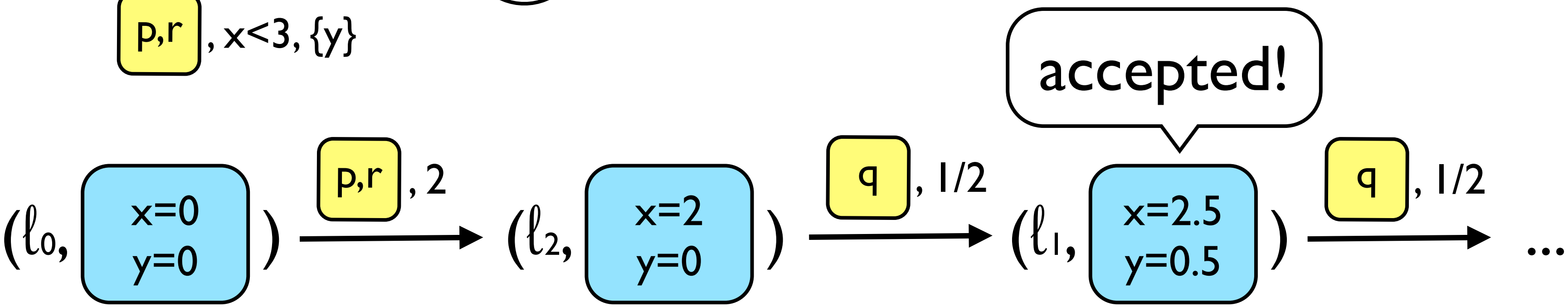
Clocks = {x,y}



Clock Guards

$$g := x \bowtie q \mid g \wedge g$$

for $\bowtie \in \{<, \leq, >, \geq\}$, $q \in \mathbb{Q}$



TA distance

(max error w.r.t. timed regular properties)

set of timed paths accepted by $\mathcal{A}$

$$TA(s,s') = \sup_{\mathcal{A} \in TA} |P[s](\{\pi \in L(\mathcal{A})\}) - P[s'](\{\pi \in L(\mathcal{A})\})|$$

TA distance

(max error w.r.t. timed regular properties)

set of timed paths accepted by $\mathcal{A}$

measurable in $\sigma(\mathcal{T})$

$$TA(s,s') = \sup_{\mathcal{A} \in TA} |P[s](\{\pi \in L(\mathcal{A})\}) - P[s'](\{\pi \in L(\mathcal{A})\})|$$

Relation with Trace Distance

$$TA(s,s') \leq d(s,s') = \sup_{E \in \sigma(\mathcal{T})} |P[s](E) - P[s'](E)|$$

TA distance

(max error w.r.t. timed regular properties)

set of timed paths accepted by $\mathcal{A}$

measurable in $\sigma(\mathcal{T})$

$$TA(s,s') = \sup_{\mathcal{A} \in TA} |P[s](\{\pi \in L(\mathcal{A})\}) - P[s'](\{\pi \in L(\mathcal{A})\})|$$

Relation with Trace Distance

$$TA(s,s') = d(s,s') = \sup_{E \in \sigma(\mathcal{T})} |P[s](E) - P[s'](E)|$$

The theorem behind...

For $\mu,\nu\colon \Sigma \to \mathbb{R}_+$ finite measures on (X,Σ)

and $F \subseteq \Sigma$ field such that $\sigma(F)=\Sigma$

Representation Theorem

$$\|\mu - \nu\| = \sup_{E \in F} |\mu(E) - \nu(E)|$$

The theorem behind...

For $\mu,\nu: \Sigma \to \mathbb{R}_+$ finite measures on (X,Σ)
and $F \subseteq \Sigma$ field such that $\sigma(F)=\Sigma$

Representation Theorem

$$\|\mu - \nu\| = \sup_{E \in F} |\mu(E) - \nu(E)|$$

F is much simpler than Σ, nevertheless
it suffices to attain the supremum!

A series of characterizations

$$\mathrm{MTL}(s,s') = \mathrm{MTL}^{\neg U}(s,s')$$

$$\mathrm{TA}(s,s') = \mathrm{DTA}(s,s') = 1\text{-}\mathrm{DTA}(s,s') = 1\text{-}\mathrm{RDTA}(s,s')$$

A series of characterizations

max error w.r.t. $\varphi \in MTL$
without Until

$$MTL(s,s') = MTL^{\neg U}(s,s')$$

$$TA(s,s') = DTA(s,s') = 1\text{-}DTA(s,s') = 1\text{-}RDTA(s,s')$$

A series of characterizations

max error w.r.t. $\varphi \in$ MTL without Until

$$MTL(s,s') = MTL^{\neg U}(s,s')$$

max error w.r.t. Deterministic TAs

$$TA(s,s') = DTA(s,s') = 1\text{-}DTA(s,s') = 1\text{-}RDTA(s,s')$$

A series of characterizations

max error w.r.t. φ∈MTL without Until

$$MTL(s,s') = MTL^{\neg U}(s,s')$$

max error w.r.t. Deterministic TAs

$$TA(s,s') = DTA(s,s') = 1\text{-}DTA(s,s') = 1\text{-}RDTA(s,s')$$

max error w.r.t. single-clock DTAs

A series of characterizations

max error w.r.t. φ∈MTL without Until

$$\mathrm{MTL}(s,s') = \mathrm{MTL}^{\neg U}(s,s')$$

max error w.r.t. Deterministic TAs

max error w.r.t. Resetting 1-DTAs

$$\mathrm{TA}(s,s') = \mathrm{DTA}(s,s') = \text{1-DTA}(s,s') = \text{1-RDTA}(s,s')$$

max error w.r.t. single-clock DTAs

Computing the trace distance...

NP-hardness [Lyngsø-Pedersen JCSS'02]

Approximating the trace distance up to any ε>0 whose size is polynomial in the size of the MC is NP-hard.

Computing the trace distance...

NP-hardness [Lyngsø-Pedersen JCSS'02]

Approximating the trace distance up to any ε>0 whose size is polynomial in the size of the MC is NP-hard.

reduction from the max-clique problem

Computing the trace distance...

NP-hardness [Lyngsø-Pedersen JCSS'02]

Approximating the trace distance up to any ε>0 whose size is polynomial in the size of the MC is NP-hard.

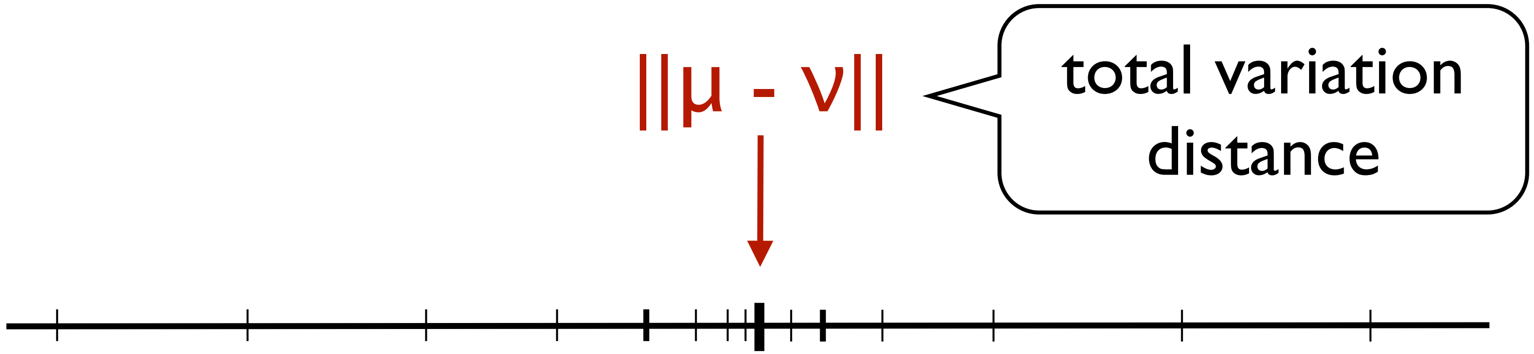
reduction from the max-clique problem

Decidability still an open problem!

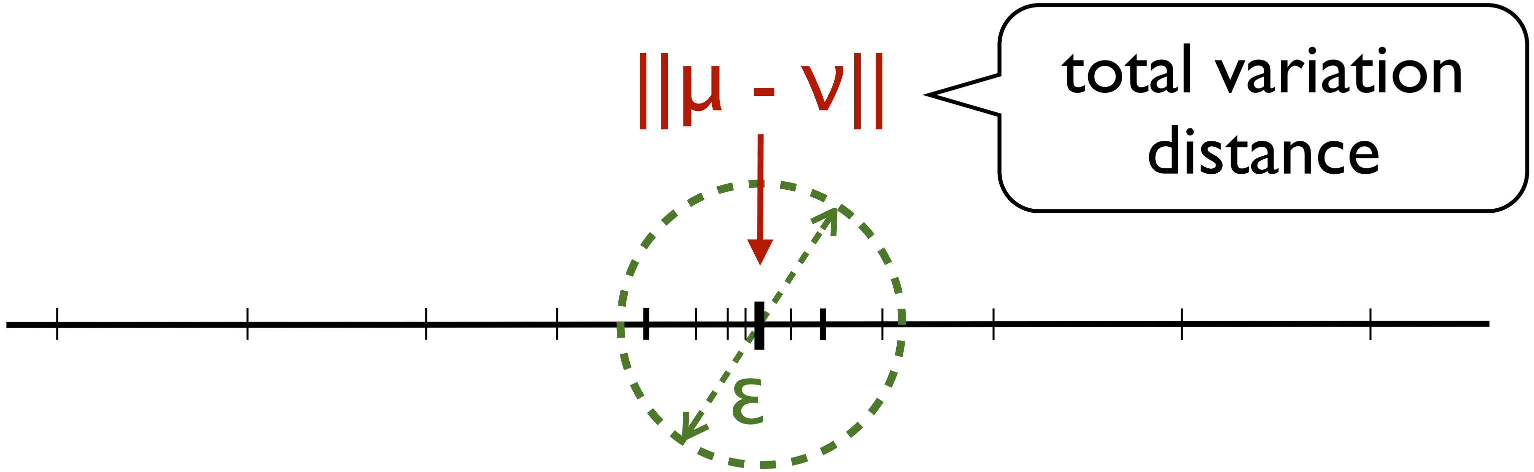
Approximation Algorithm for the Trace Distance

(from below & from above)

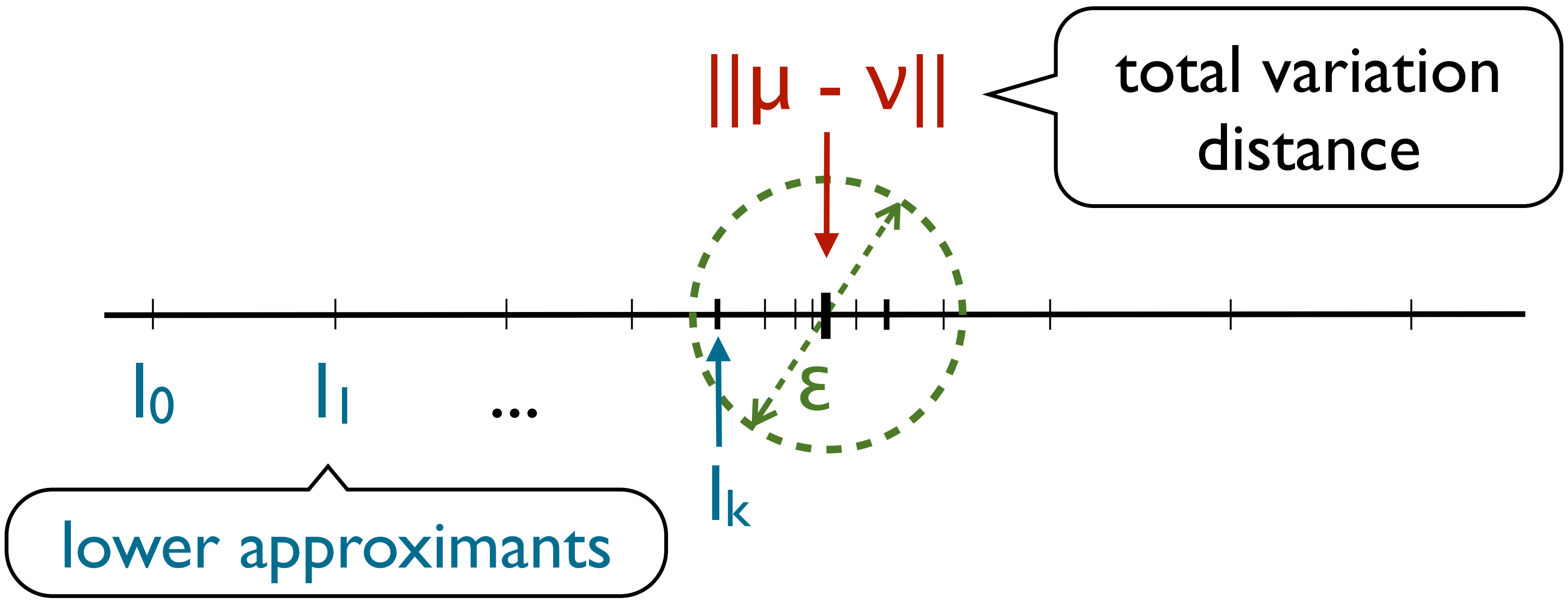
Approximation Algorithm



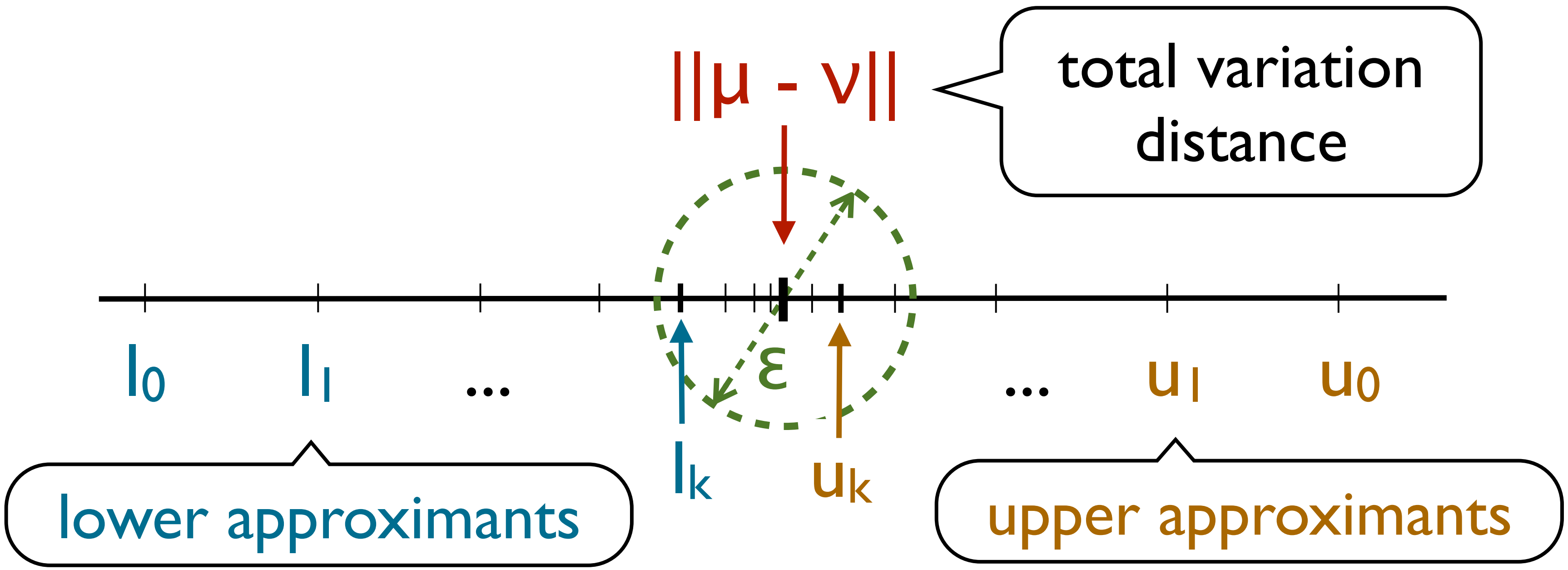
Approximation Algorithm



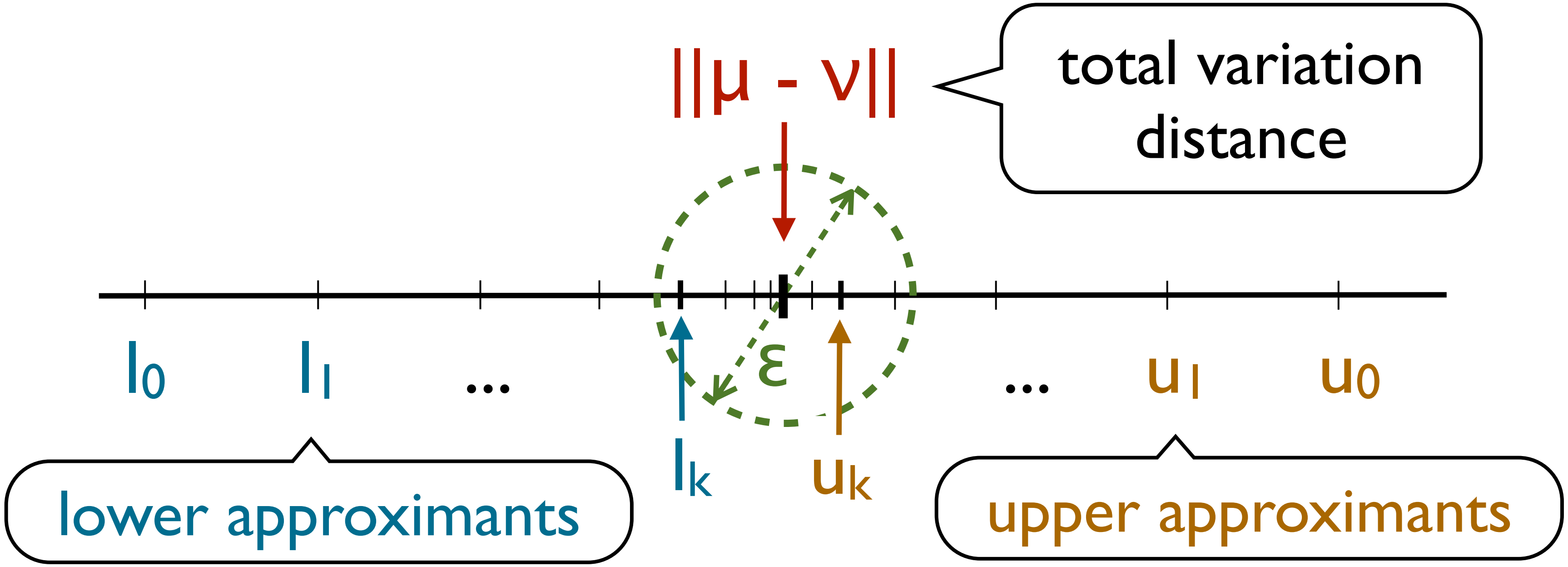
Approximation Algorithm



Approximation Algorithm

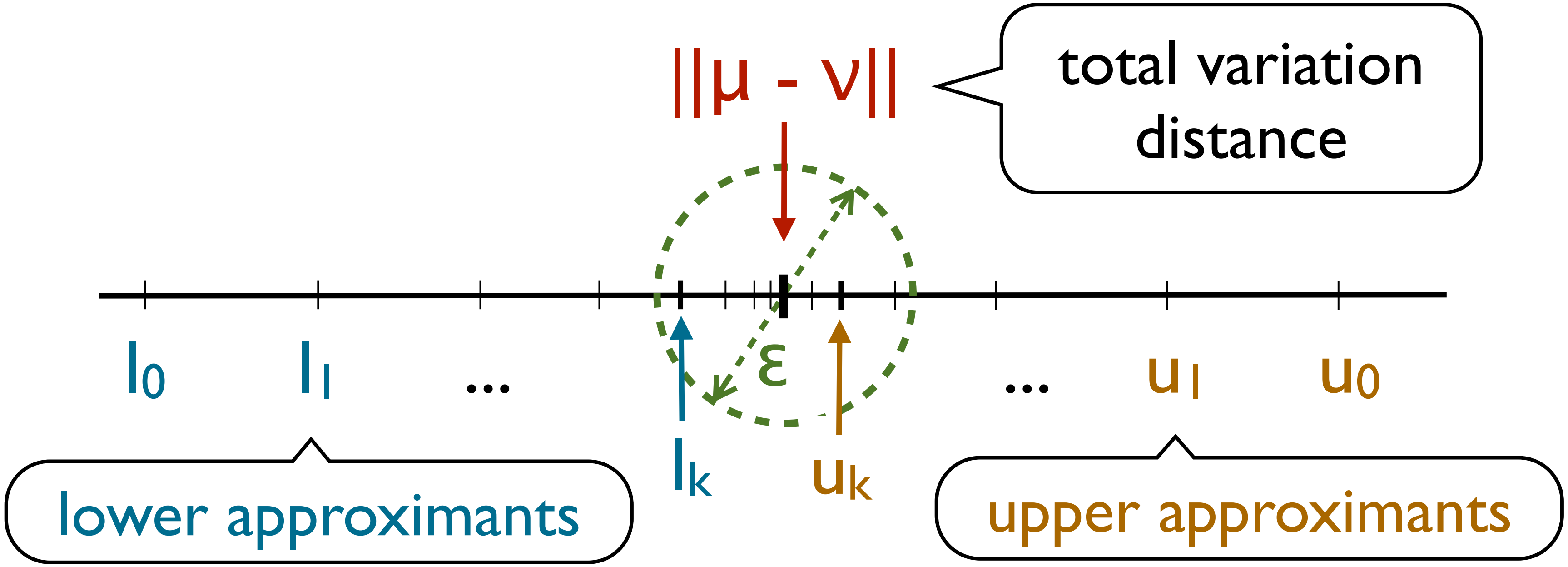


Approximation Algorithm



- l_i and u_i must converge to $||\mu - \nu||$,

Approximation Algorithm



- l_i and u_i must converge to $||\mu - \nu||$,
- For all $i \in \mathbb{N}$, l_i and u_i must be *computable*.

... from below

... from below

Representation Theorem

recall that...

$$\|\mu - \nu\| = \sup_{E \in F} |\mu(E) - \nu(E)|$$

... from below

Representation Theorem

recall that...

$$||\mu - \nu|| = \sup_{E \in F} |\mu(E) - \nu(E)|$$

F field that generates Σ

... from below

Representation Theorem

recall that...

$$\|\mu - \nu\| = \sup_{E \in F} |\mu(E) - \nu(E)|$$

F field that generates Σ

We need $F_0 \subseteq F_1 \subseteq F_2 \subseteq \ldots$ such that $\bigcup_i F_i = F$

$$l_i = \sup_{E \in F_i} |\mu(E) - \nu(E)|$$

... from below

Representation Theorem — recall that...

$$||\mu - \nu|| = \sup_{E \in F} |\mu(E) - \nu(E)|$$

F field that generates Σ

We need $F_0 \subseteq F_1 \subseteq F_2 \subseteq \ldots$ such that $\bigcup_i F_i = F$

$$l_i = \sup_{E \in F_i} |\mu(E) - \nu(E)|$$

so that $\forall i \geq 0,\ l_i \leq l_{i+1}$ & $\sup_i l_i = ||\mu - \nu||$

increasing — limiting

Approx. Trace Distance

from below

Provide $F_0 \subseteq F_1 \subseteq F_2 \subseteq \ldots$ such that

$\bigcup_i F_i$ is a field for $\sigma(\mathcal{T})$

Approx. Trace Distance

from below

Provide $F_0 \subseteq F_1 \subseteq F_2 \subseteq \ldots$ such that
$\cup_i F_i$ is a field for $\sigma(\mathcal{T})$

Take F_i to be the collection of finite unions of cylinders

$$\mathfrak{C}(L_0, R_0, \ldots, R_{i-1}, L_i) \in \mathcal{T}$$

where $R_j \in \{[\frac{n}{2^i}, \frac{n+1}{2^i}) \mid 0 \leq n \leq i2^i\} \cup \{[i, \infty)\}$

Approx. Trace Distance

from below

Provide $F_0 \subseteq F_1 \subseteq F_2 \subseteq \ldots$ such that

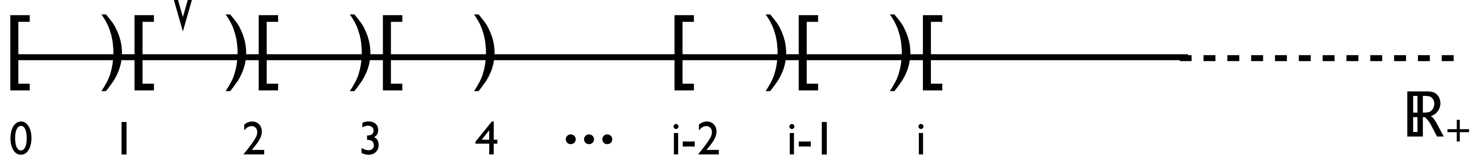
$\cup_i F_i$ is a field for $\sigma(\mathcal{T})$

Take F_i to be the collection of finite unions of cylinders

$$\mathfrak{C}(L_0, R_0, \ldots, R_{i-1}, L_i) \in \mathcal{T}$$

where $R_j \in \{[\frac{n}{2^i}, \frac{n+1}{2^i}) \mid 0 \leq n \leq i2^i\} \cup \{[i, \infty)\}$

each repartitioned in 2^i [closed-open) intervals



... from above

... from above

Coupling Characterization

$$||\mu - \nu|| = \min \{w(\neq) \mid w \in \Omega(\mu,\nu)\}$$

it is know that...

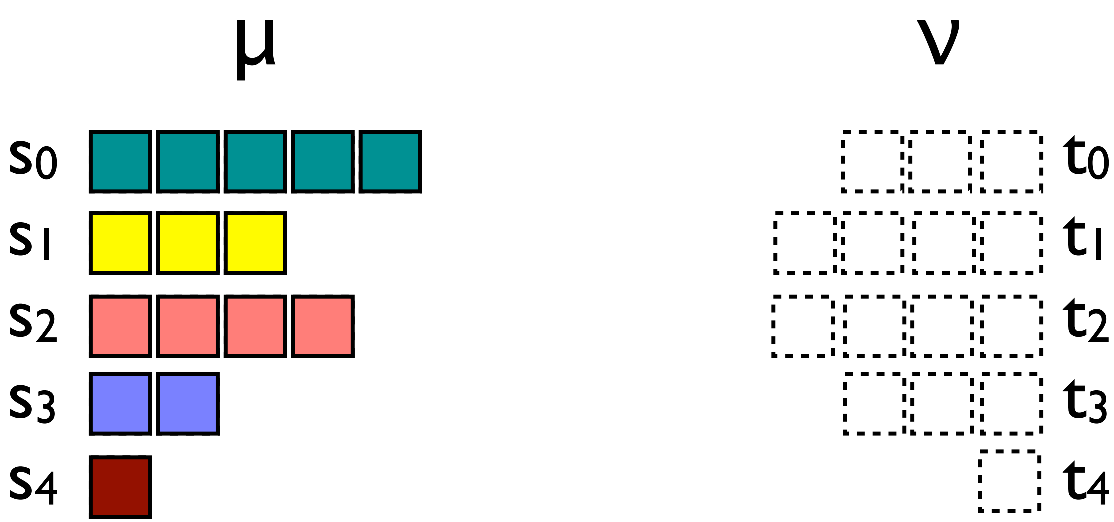
... from above

it is know that...

Coupling Characterization

$$\|\mu - \nu\| = \min \{ w(\neq) \mid w \in \Omega(\mu,\nu) \}$$

Coupling as a transportation schedule...



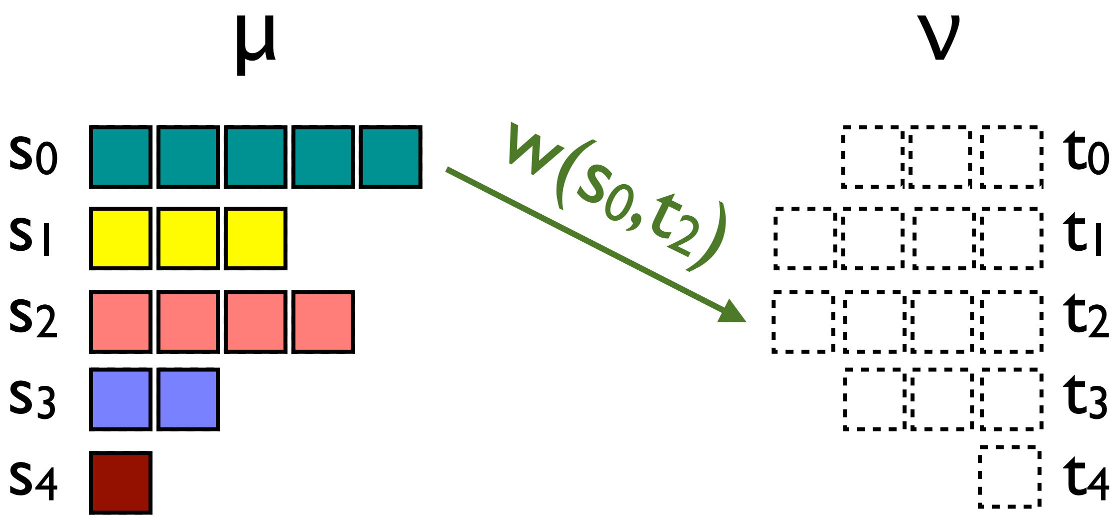
... from above

Coupling Characterization

$$||\mu - \nu|| = \min \{w(\neq) \mid w \in \Omega(\mu,\nu)\}$$

it is know that...

Coupling as a transportation schedule...



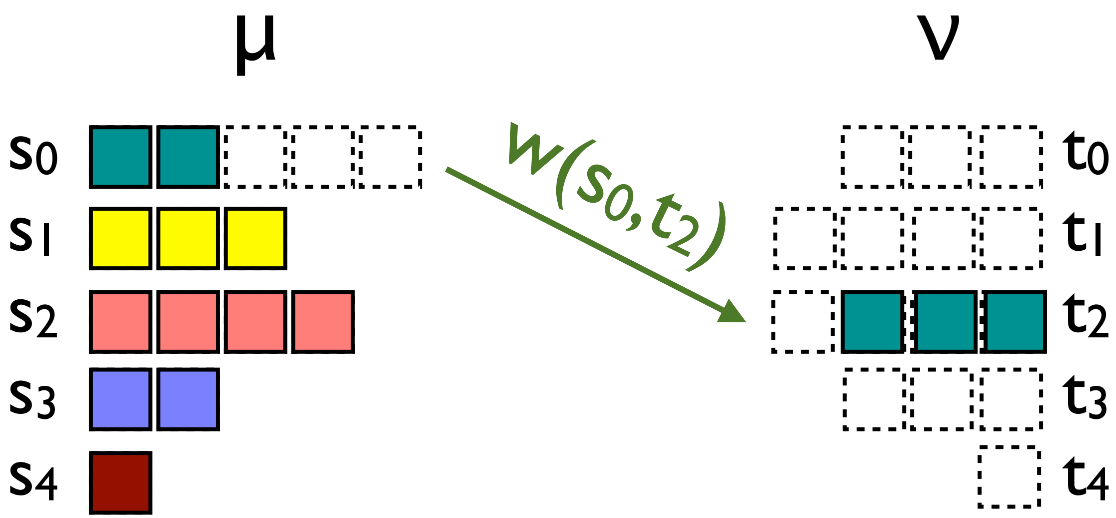
... from above

Coupling Characterization

it is know that...

$$||\mu - \nu|| = \min \{w(\neq) \mid w \in \Omega(\mu,\nu)\}$$

Coupling as a transportation schedule...



... from above

Coupling Characterization

$$||\mu - \nu|| = \min \{w(\neq) \mid w \in \Omega(\mu,\nu)\}$$

it is know that...

... from above

Coupling Characterization

$$||\mu - \nu|| = \min \{w(\neq) \mid w \in \Omega(\mu,\nu)\}$$

it is know that...

We need $\Omega_0 \subseteq \Omega_1 \subseteq \Omega_2 \subseteq \ldots$ such that

$\bigcup_i \Omega_i$ dense in $\Omega(\mu,\nu)$ w.r.t. total variation

$$u_i = \inf \{w(\neq) \mid w \in \Omega_i\}$$

... from above

it is know that...

Coupling Characterization

$$||\mu - \nu|| = \min \{w(\neq) \mid w \in \Omega(\mu,\nu)\}$$

We need $\Omega_0 \subseteq \Omega_1 \subseteq \Omega_2 \subseteq \ldots$ such that

$\cup_i \Omega_i$ dense in $\Omega(\mu,\nu)$ w.r.t. total variation

$$u_i = \inf \{w(\neq) \mid w \in \Omega_i\}$$

so that $\forall i \geq 0, u_i \geq u_{i+1}$ & $\inf_i u_i = ||\mu - \nu||$

decreasing

limiting

Approx. Trace Distance

from above

Provide $\Omega_0 \subseteq \Omega_1 \subseteq \Omega_2 \subseteq \ldots$ such that

$\bigcup_i \Omega_i$ is dense in $\Omega(P[s],P[s'])$

Approx. Trace Distance

from above

Provide $\Omega_0 \subseteq \Omega_1 \subseteq \Omega_2 \subseteq \ldots$ such that

$\bigcup_i \Omega_i$ is dense in $\Omega(P[s],P[s'])$

Take $\Omega_i = \{P_{\mathcal{C}}[s,s'] \in \Omega(P[s],P[s']) \mid \mathcal{C}$ of rank $2^i\}$

where $P_{\mathcal{C}}[s,s']$ is the probability generated by

Approx. Trace Distance

from above

Provide $\Omega_0 \subseteq \Omega_1 \subseteq \Omega_2 \subseteq \ldots$ such that

$\bigcup_i \Omega_i$ is dense in $\Omega(P[s],P[s'])$

Take $\Omega_i = \{P_{\mathcal{C}}[s,s'] \in \Omega(P[s],P[s']) \mid \mathcal{C} \text{ of rank } 2^i\}$

where $P_{\mathcal{C}}[s,s']$ is the probability generated by

coupling structure of rank k

$\mathcal{C}: S \times S \to \Delta(\Pi^k S \times \Pi^k S)$

such that $\mathcal{C}(s,s') \in \Omega(P[s]^k, P[s']^k)$

Stochastic process generating pairs of timed paths divided in multisteps of length k

Decidability

- A1: residence-time distributions are computable on [q,q') with $q,q' \in \mathbb{Q}_+$
- A2: total variation between residence-time distributions is computable

For any ε>0, the approximation procedure for the trace distance is decidable.

Decidability

Not that strong!

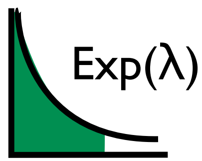
- A1: residence-time distributions are computable on $[q,q')$ with $q,q' \in \mathbb{Q}_+$
- A2: total variation between residence-time distributions is computable

For any ε>0, the approximation procedure for the trace distance is decidable.

Decidability

Not that strong!

- A1: residence-time distributions are computable on [q,q') with $q,q' \in \mathbb{Q}_+$
- A2: total variation between residence-time distributions is computable

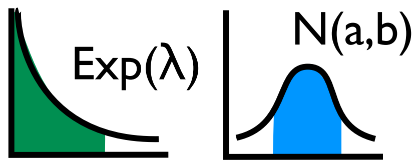


For any ε>0, the approximation procedure for the trace distance is decidable.

Decidability

Not that strong!

- A1: residence-time distributions are computable on [q,q') with $q,q' \in \mathbb{Q}_+$
- A2: total variation between residence-time distributions is computable

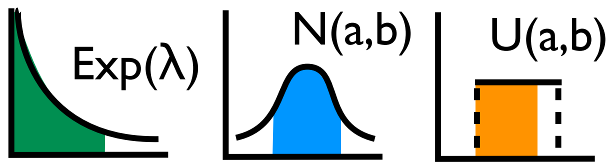


For any ε>0, the approximation procedure for the trace distance is decidable.

Decidability

Not that strong!

- A1: residence-time distributions are computable on [q,q') with $q,q' \in \mathbb{Q}_+$
- A2: total variation between residence-time distributions is computable

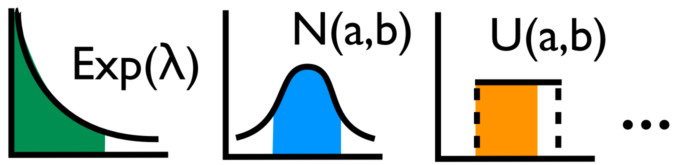


For any ε>0, the approximation procedure for the trace distance is decidable.

Decidability

Not that strong!

- A1: residence-time distributions are computable on [q,q') with $q,q' \in \mathbb{Q}_+$
- A2: total variation between residence-time distributions is computable



For any ε>0, the approximation procedure for the trace distance is decidable.

Concluding Remarks

Concluding Remarks

- Trace Distance vs Model Checking

Concluding Remarks

- Trace Distance vs Model Checking
- General results for Total Variation distance:

Concluding Remarks

- Trace Distance vs Model Checking
- General results for Total Variation distance:
    - algebraic representation theorem

Concluding Remarks

- Trace Distance vs Model Checking
- General results for Total Variation distance:
  - algebraic representation theorem
  - approximation strategies

Concluding Remarks

- Trace Distance vs Model Checking
- General results for Total Variation distance:
  - algebraic representation theorem
  - approximation strategies
- Approx. algorithm for Trace Distance

Concluding Remarks

- Trace Distance vs Model Checking
- General results for Total Variation distance:
  - algebraic representation theorem
  - approximation strategies
- Approx. algorithm for Trace Distance
- Relation with Kantorovich dist. (not shown)

Thank you

for the attention