On the Total Variation Distance of SMPs

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

25-26 September 2014

IDEA4CPS

Outline

- Semi-Markov Processes (SMPs)
- Total Variation Distance of SMPs
- Total Variation vs. Model Checking
- An Approximation Algorithm
- Concluding Remarks

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...

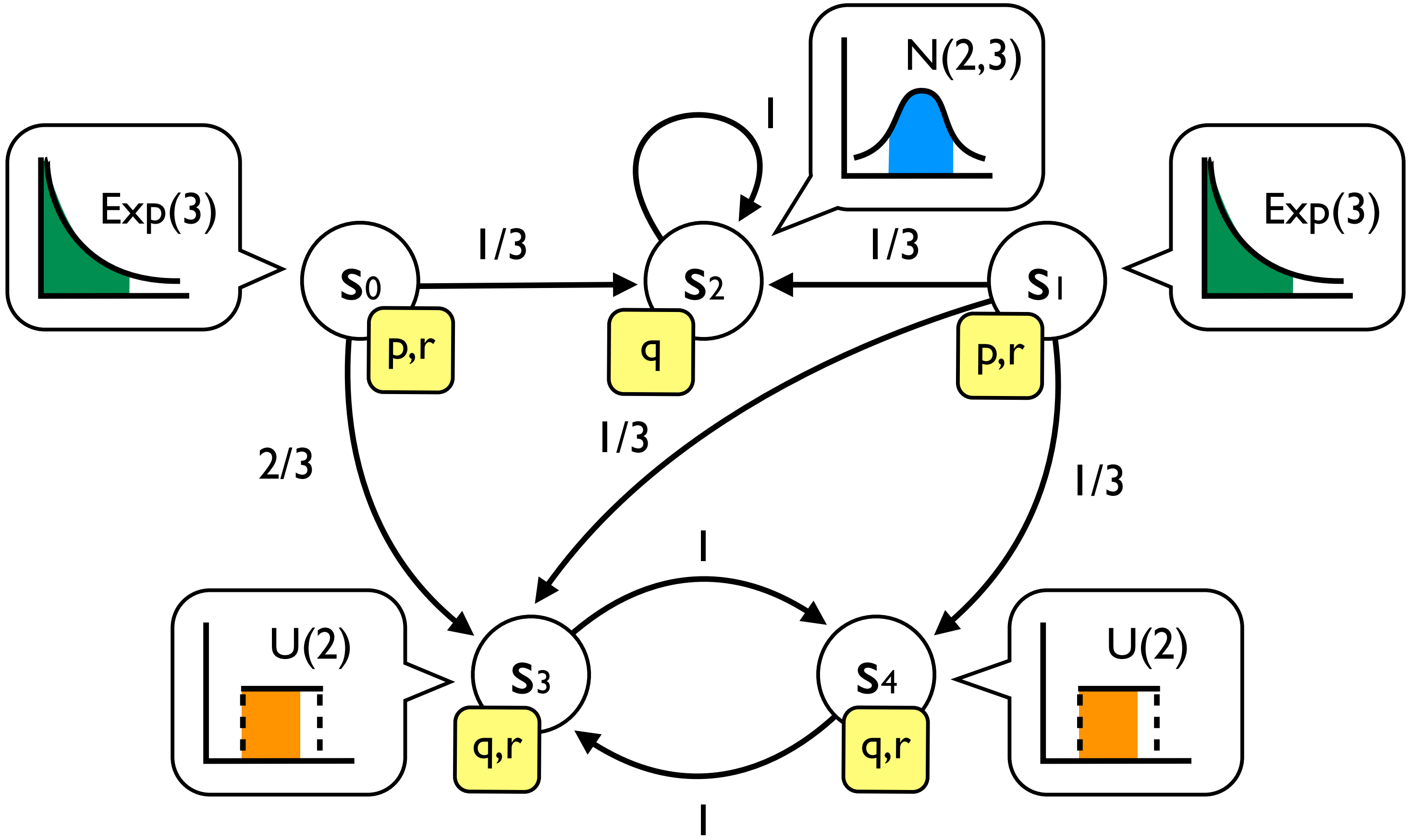
Given $\mu,\nu: \Sigma \to \mathbb{R}_+$ measures on (X,Σ)

Total Variation Distance

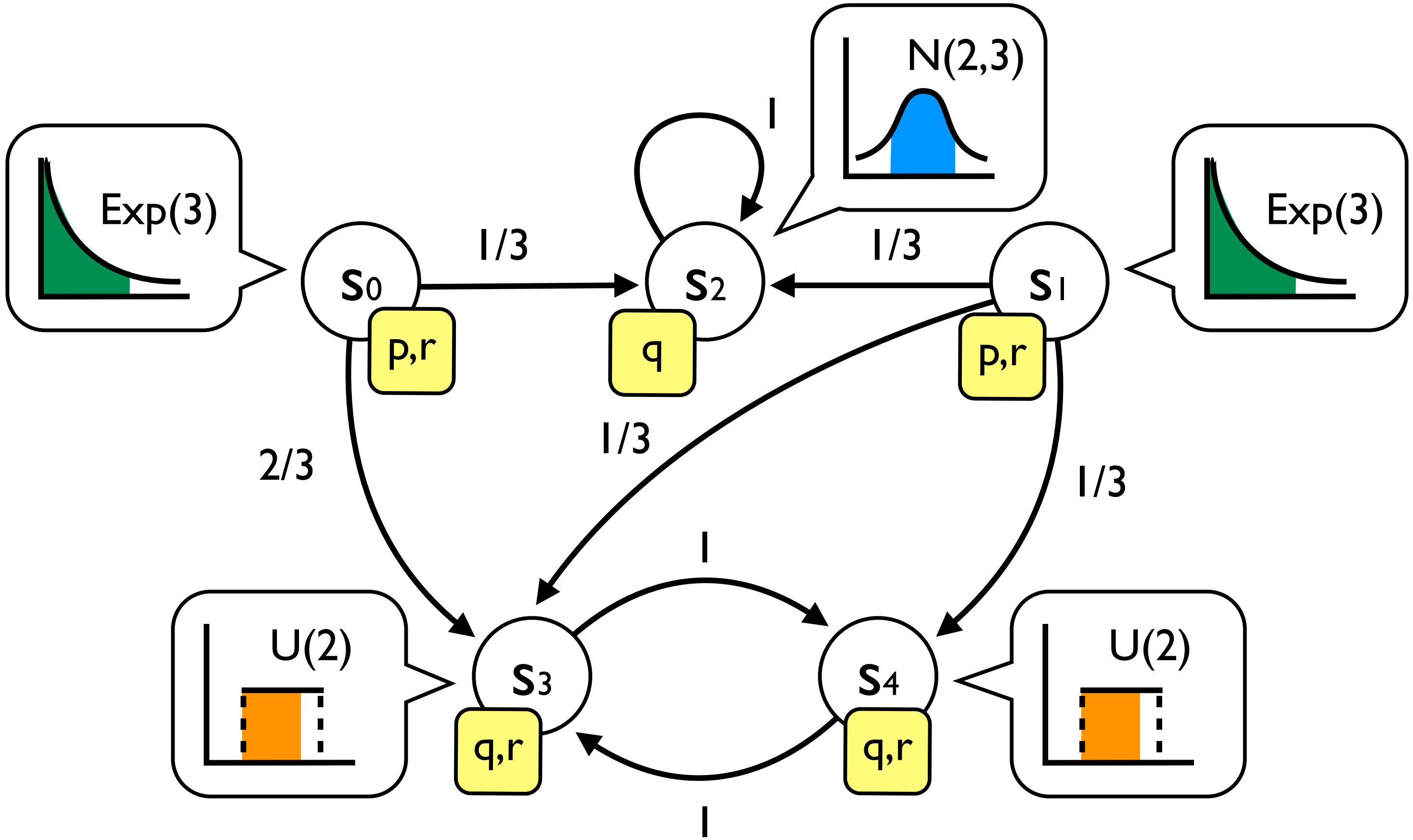
$$\| \mu - \nu \| = \sup_{E \in \Sigma} |\mu(E) - \nu(E)|$$

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

semi-Markov Processes

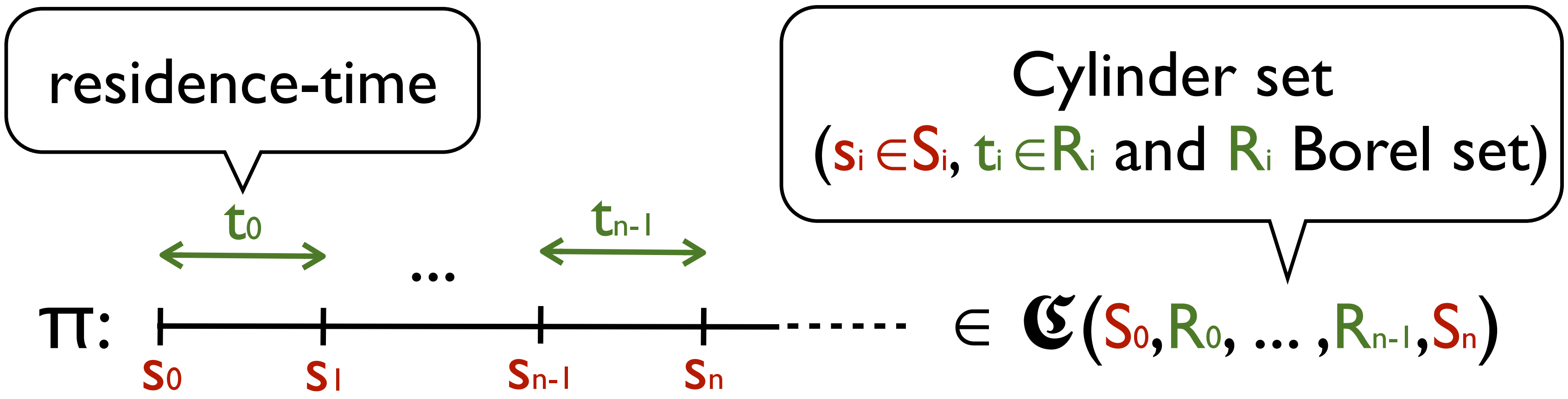


semi-Markov Processes



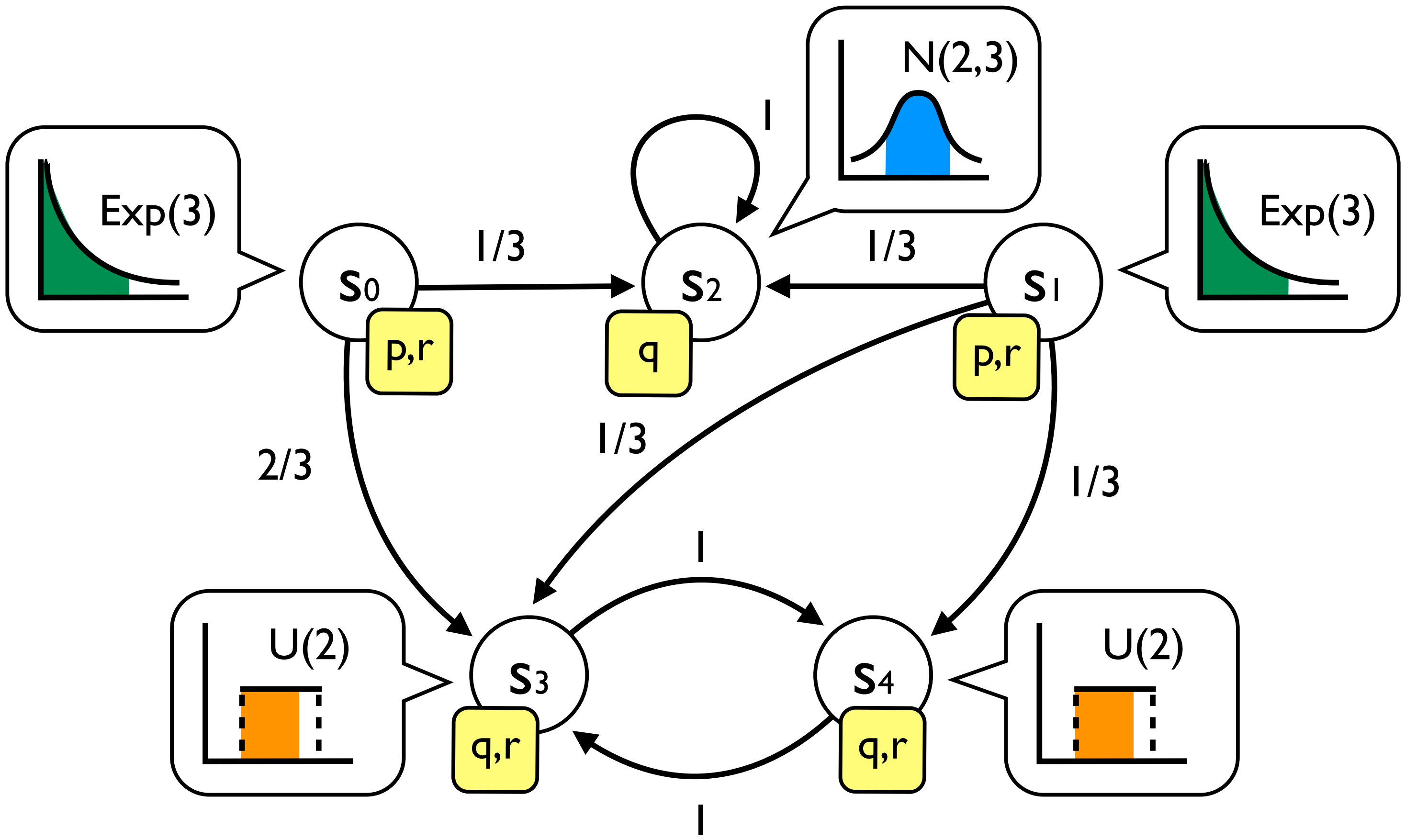
Given an initial state, SMPs can be interpreted as “machines” that emit timed traces of states at a certain probability

Timed paths & Events

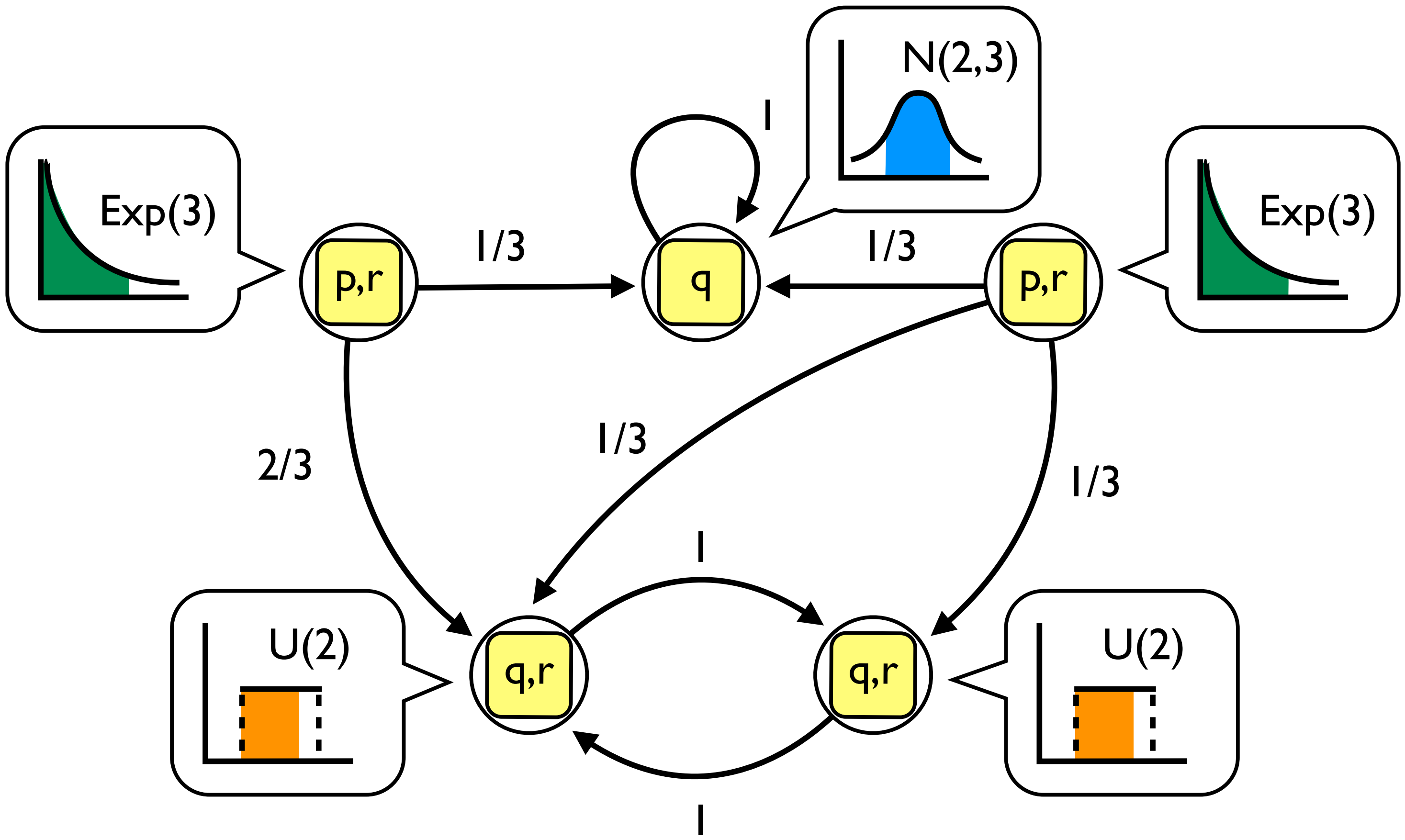


$P[s](\mathfrak{C}(S_0,R_0, \ldots ,R_{n-1},S_n))$ = "probability that, *starting from s*, the SMP emits a timed path with prefix in $S_0 \times R_0 \times \ldots \times R_{n-1} \times S_n$"

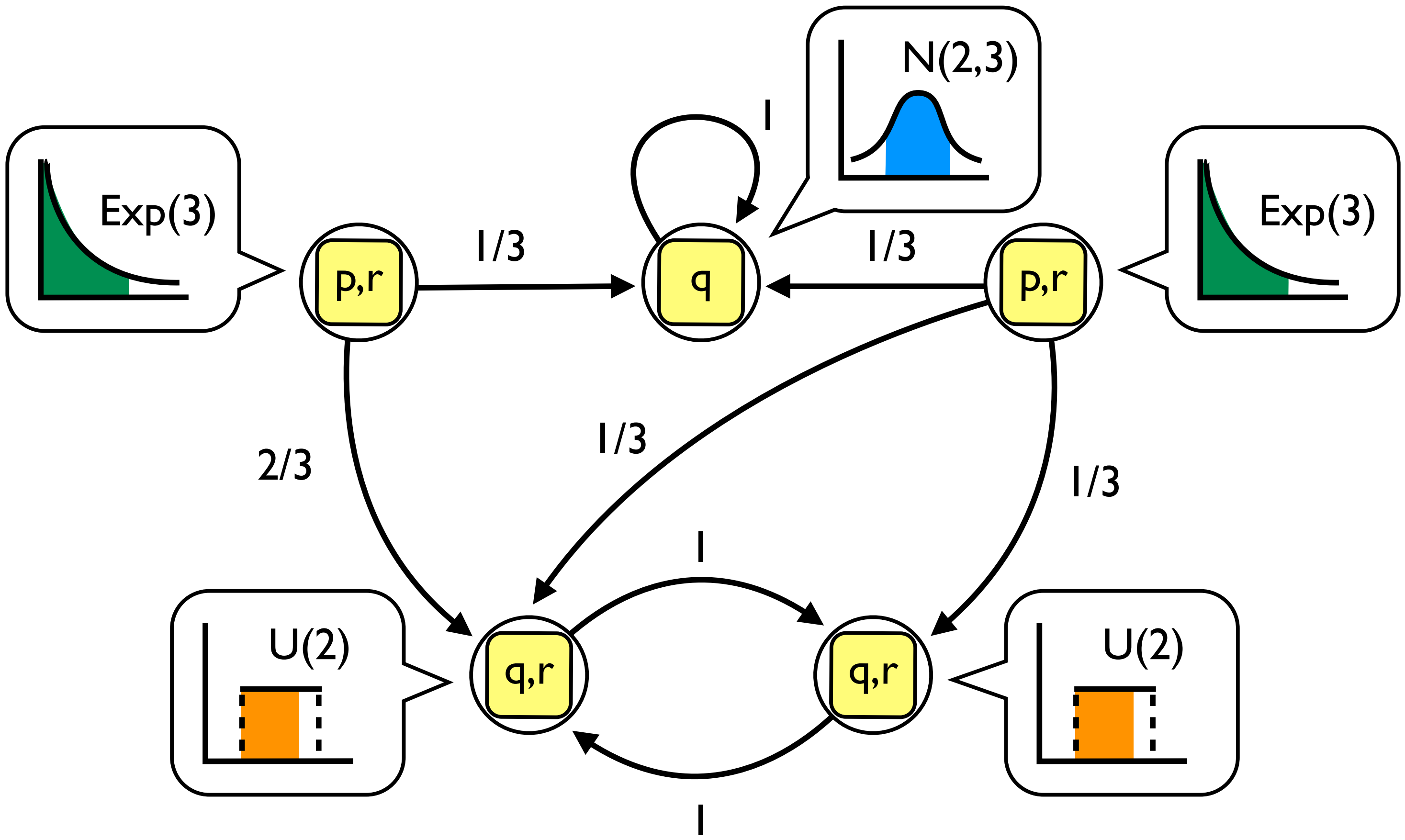
Prob. Trace Equivalence



Prob. Trace Equivalence

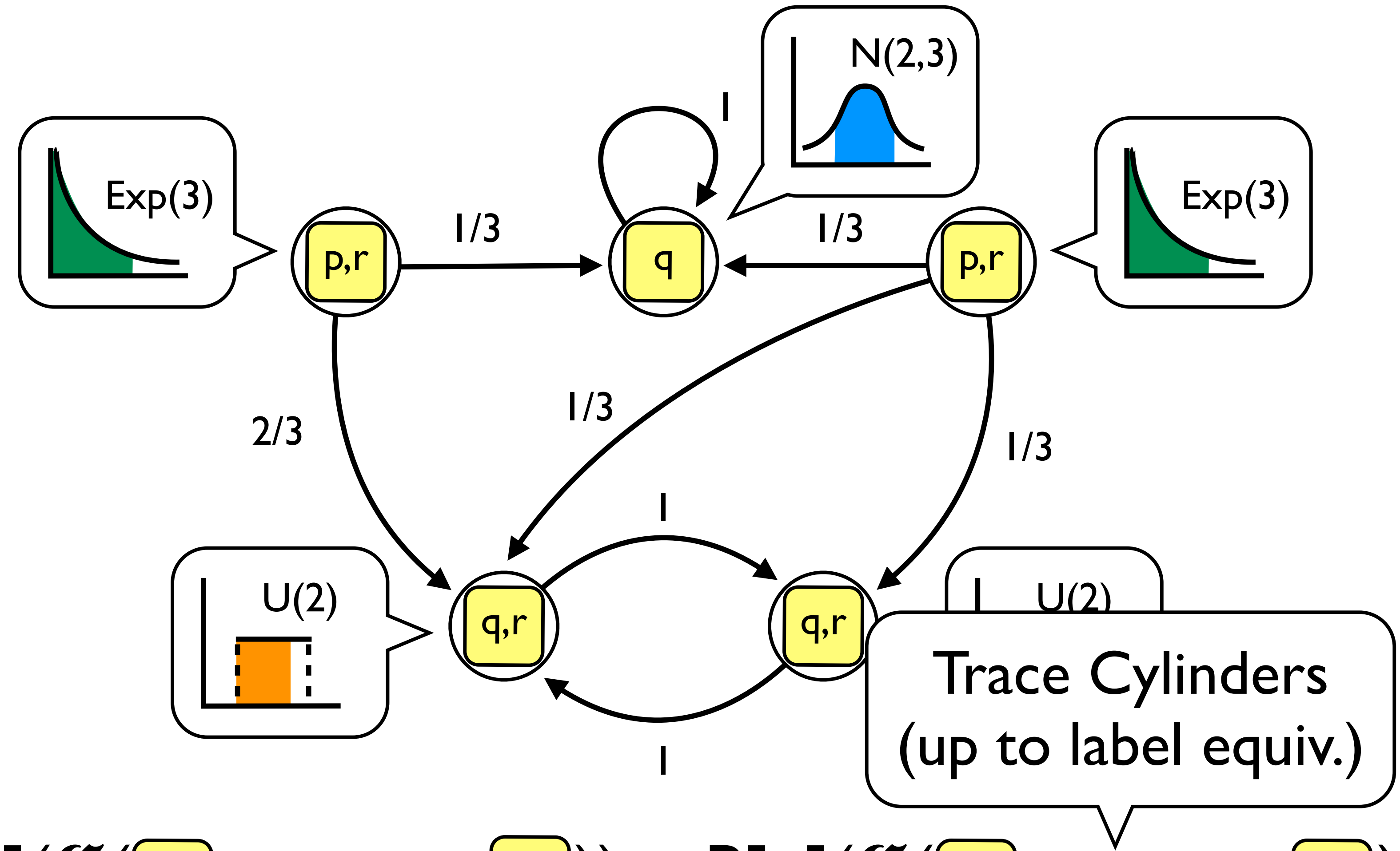


Prob. Trace Equivalence



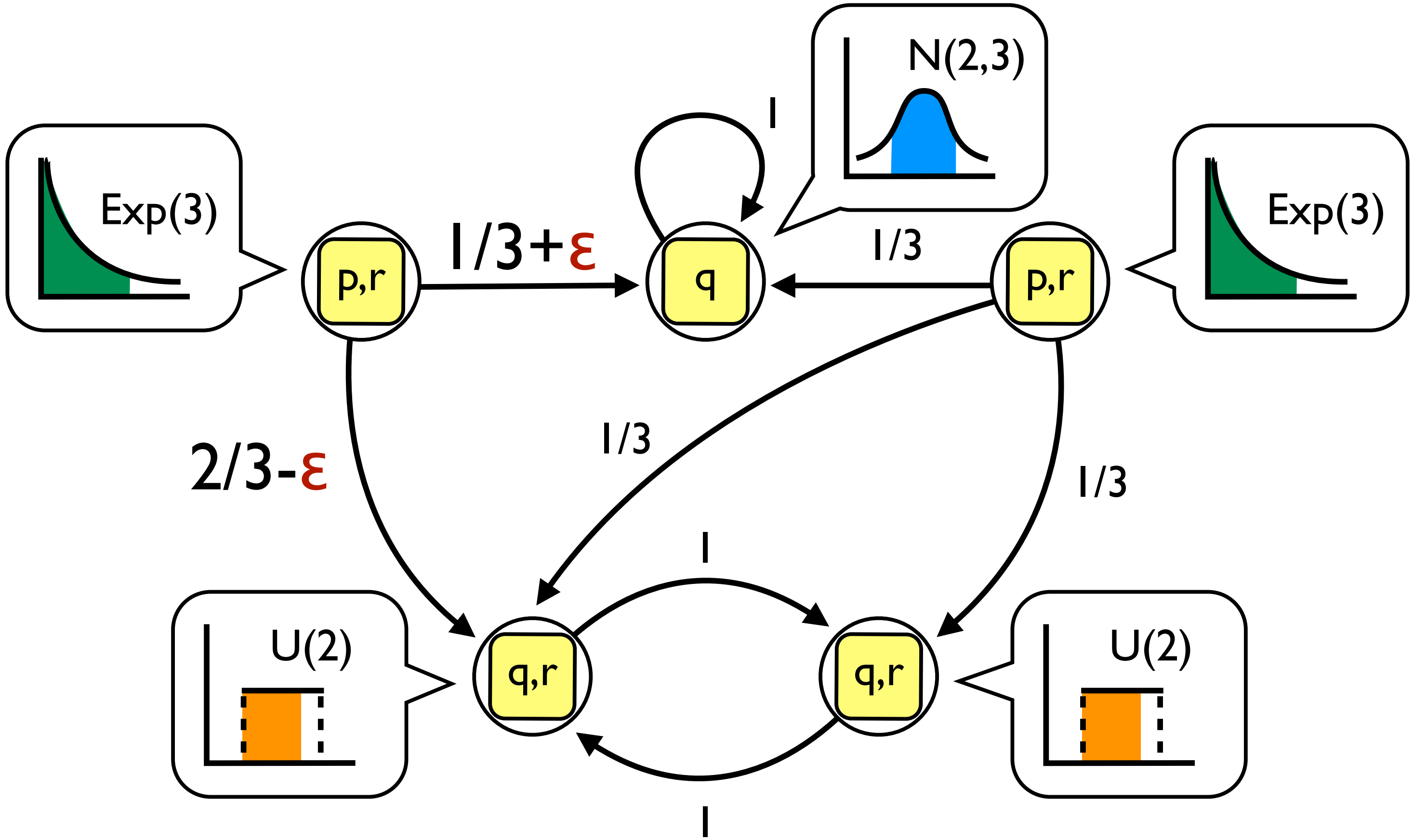
$$P[s_0](\mathfrak{C}(L_0, R_0, \ldots, R_{n-1}, L_n)) = P[s_1](\mathfrak{C}(L_0, R_0, \ldots, R_{n-1}, L_n))$$

Prob. Trace Equivalence



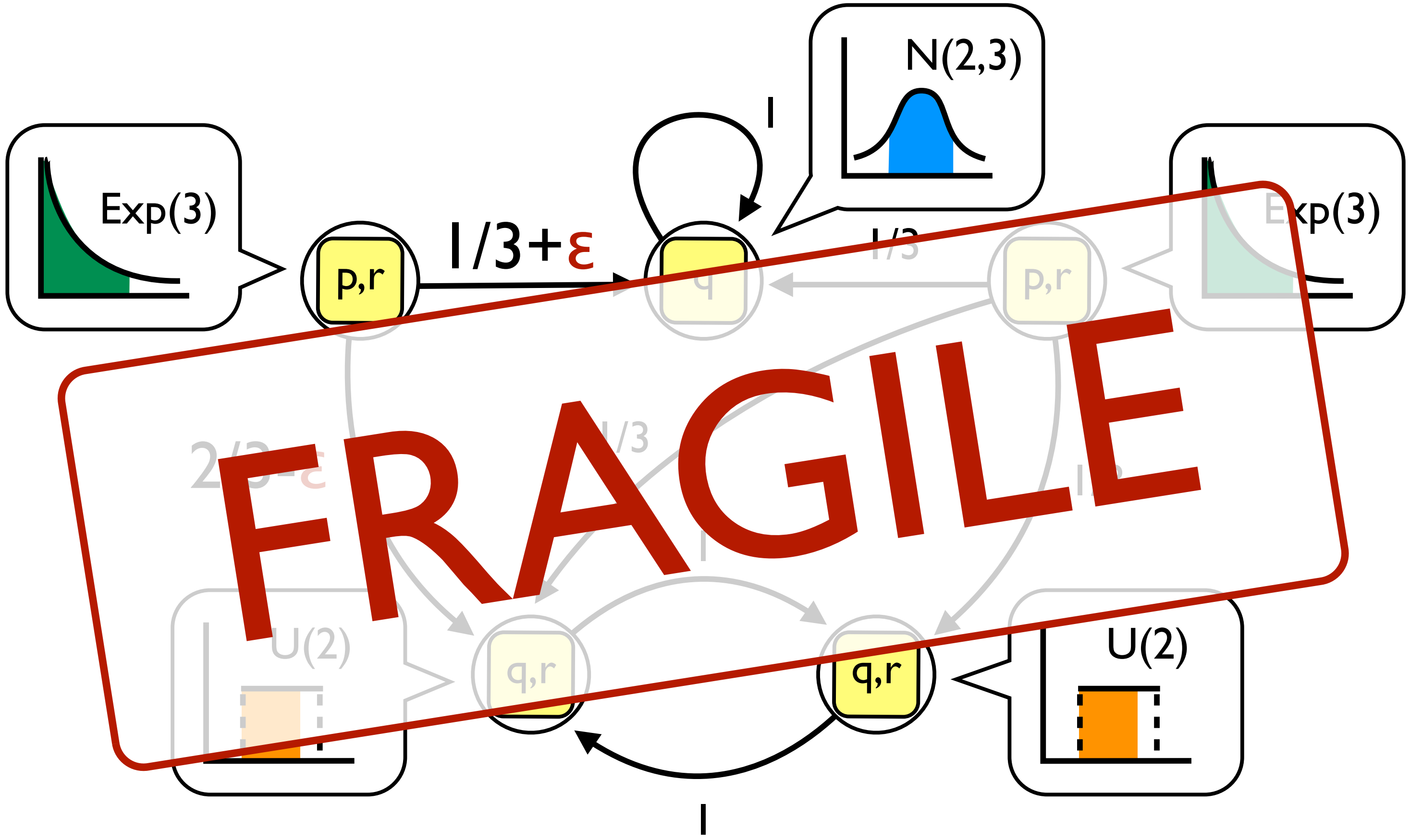
$$P[s_0](\mathfrak{C}(L_0, R_0, \ldots, R_{n-1}, L_n)) = P[s_1](\mathfrak{C}(L_0, R_0, \ldots, R_{n-1}, L_n))$$

Prob. Trace Equivalence



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

Prob. Trace Equivalence



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

Trace Pseudometric

(difference w.r.t. linear real-time behaviors)

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

σ-algebra generated from Trace Cylinders

Trace Pseudometric

(difference w.r.t. linear real-time behaviors)

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

σ-algebra generated from Trace Cylinders

It's a Behavioral Distance!

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

Trace Distance

vs.

Model Checking

(i.e., what do they have in common?)

Model Checking SMPs

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

SMP $\models$ Linear Real-time Spec.

Model Checking SMPs

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

a proper measurable set!

SMP $\models$ Linear Real-time Spec.

Model Checking SMPs

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

a proper measurable set!

SMP $\models$ Linear Real-time Spec.

represented as
Metric Temporal Logic
formulas

Model Checking SMPs

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

a proper measurable set!

SMP $\models$ Linear Real-time Spec.

represented as
Metric Temporal Logic
formulas

... or languages
recognized
by Timed Automata

Metric Temporal Logic

(Alur-Henzinger)

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

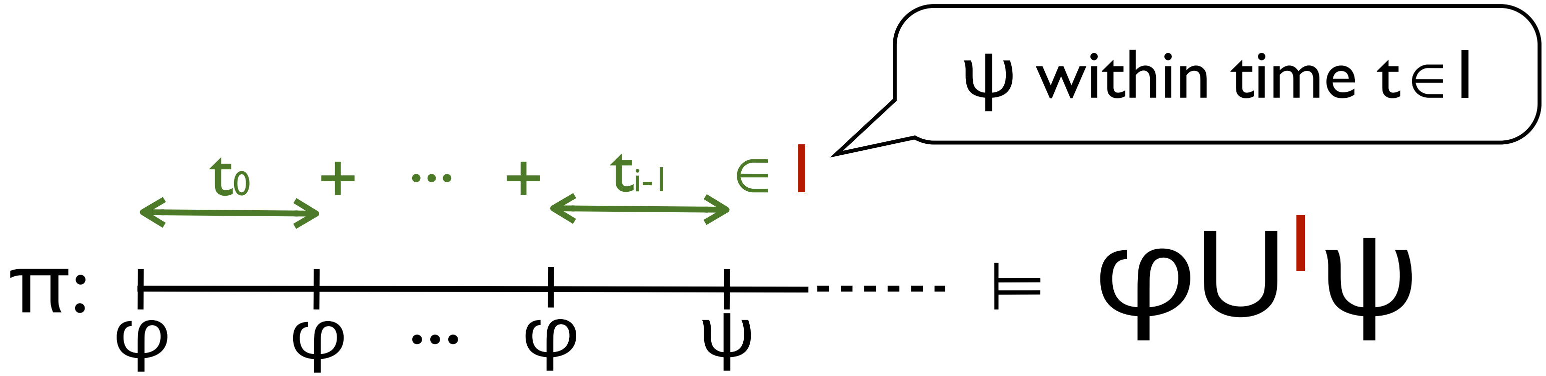
Metric Temporal Logic

(Alur-Henzinger)

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

(difference w.r.t. MTL properties)

set of timed paths that satisfy φ

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

MTL distance

(difference w.r.t. MTL properties)

set of timed paths that satisfy φ

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

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

Relation with Trace Distance

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

MTL distance

(difference w.r.t. MTL properties)

set of timed paths that satisfy φ

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

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

Relation with Trace Distance

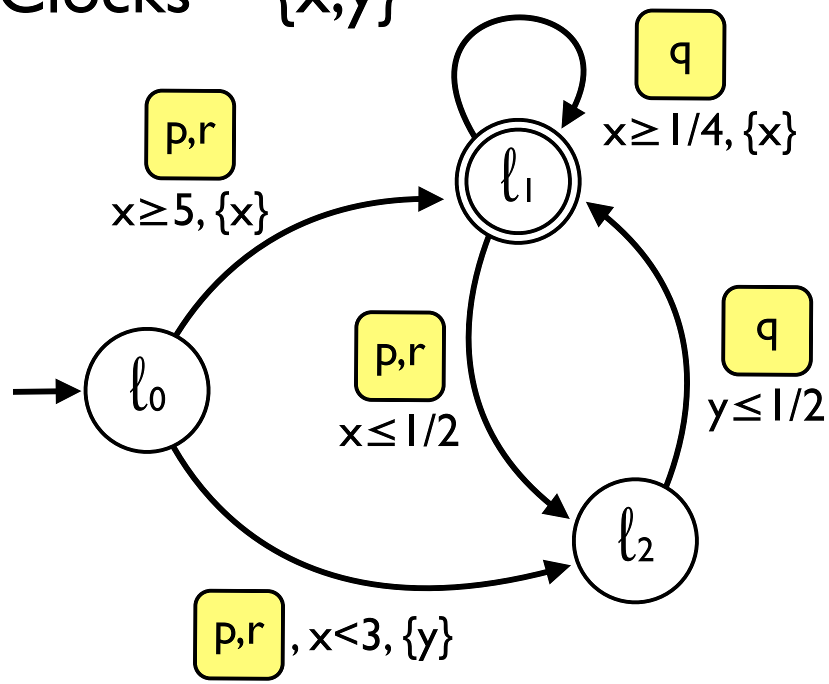
$$MTL(s,s') = d(s,s') = \sup_{E \in \sigma(\mathscr{T})} |P[s](E) - P[s'](E)|$$

Timed Automata

(Alur-Dill)

without invariants

Clocks = {x,y}



Clock Guards

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

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

accepted!

$(\ell_0, \begin{smallmatrix} x=0 \\ y=0 \end{smallmatrix}) \xrightarrow{\text{p,r},\ 2} (\ell_2, \begin{smallmatrix} x=2 \\ y=0 \end{smallmatrix}) \xrightarrow{\text{q},\ 1/2} (\ell_1, \begin{smallmatrix} x=2.5 \\ y=0.5 \end{smallmatrix}) \xrightarrow{\text{q},\ 1/2} \ldots$

TA distance

(difference w.r.t. regular TA properties)

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

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

TA distance

(difference w.r.t. regular TA properties)

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

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

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

Relation with Trace Distance

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

TA distance

(difference w.r.t. regular TA properties)

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

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

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

Relation with Trace Distance

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

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

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

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 to the supremum!

A series of characterizations

$$
d(s,s') = \begin{cases} MTL(s,s') = MTL^{\neg U}(s,s') \\ TA(s,s') = DTA(s,s') \\ 1\text{-}DTA(s,s') = 1\text{-}RDTA(s,s') \end{cases}
$$

A series of characterizations

distance w.r.t. φ∈MTL without Until

$$d(s,s') = \begin{cases} MTL(s,s') = MTL^{\neg U}(s,s') \\ TA(s,s') = DTA(s,s') \\ 1\text{-}DTA(s,s') = 1\text{-}RDTA(s,s') \end{cases}$$

A series of characterizations

distance w.r.t. $\varphi \in$ MTL without Until

distance w.r.t. only Deterministic TAs

$$d(s,s') = \begin{cases} MTL(s,s') = MTL^{\neg U}(s,s') \\ TA(s,s') = DTA(s,s') \\ 1\text{-}DTA(s,s') = 1\text{-}RDTA(s,s') \end{cases}$$

A series of characterizations

distance w.r.t. $\varphi \in$ MTL without Until

distance w.r.t. only Deterministic TAs

distance w.r.t. only single-clock DTAs

$$d(s,s') = \begin{cases} MTL(s,s') = MTL^{\neg U}(s,s') \\ TA(s,s') = DTA(s,s') \\ 1\text{-}DTA(s,s') = 1\text{-}RDTA(s,s') \end{cases}$$

A series of characterizations

$$d(s,s') = \begin{cases} MTL(s,s') = MTL^{\neg U}(s,s') \\ TA(s,s') = DTA(s,s') \\ 1\text{-}DTA(s,s') = 1\text{-}RDTA(s,s') \end{cases}$$

distance w.r.t. φ∈MTL without Until

distance w.r.t. only Deterministic TAs

distance w.r.t. only Resetting 1-DTAs

distance w.r.t. only single-clock DTAs

Approximation Algorithm for the Trace Distance

(from below & from above)

... 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

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

recall that...

F field that generates Σ

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

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

... 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 $\bigcup_i F_i = F$ to define

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

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

increasing — limiting

... from above

... from above

Alternative Characterization

$$\| \mu - \nu \| = 1 - \mu \wedge \nu(X)$$

it is know that...

... from above

Alternative Characterization

$$|| \mu - \nu || = 1 - \mu \wedge \nu(X)$$

it is know that...

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

$$u_i = 1 - \sup \{ m(X) \mid m \leq_{F_i} \mu \ \& \ m \leq_{F_i} \nu \}$$

... from above

Alternative Characterization

$$\| \mu - \nu \| = 1 - \mu \wedge \nu(X)$$

it is know that...

field that generates Σ

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

$$u_i = 1 - \sup \{ m(X) \mid m \leq_{F_i} \mu \ \& \ m \leq_{F_i} \nu \}$$

... from above

Alternative Characterization

$$\| \mu - \nu \| = 1 - \mu \wedge \nu(X)$$

it is know that...

field that generates Σ

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

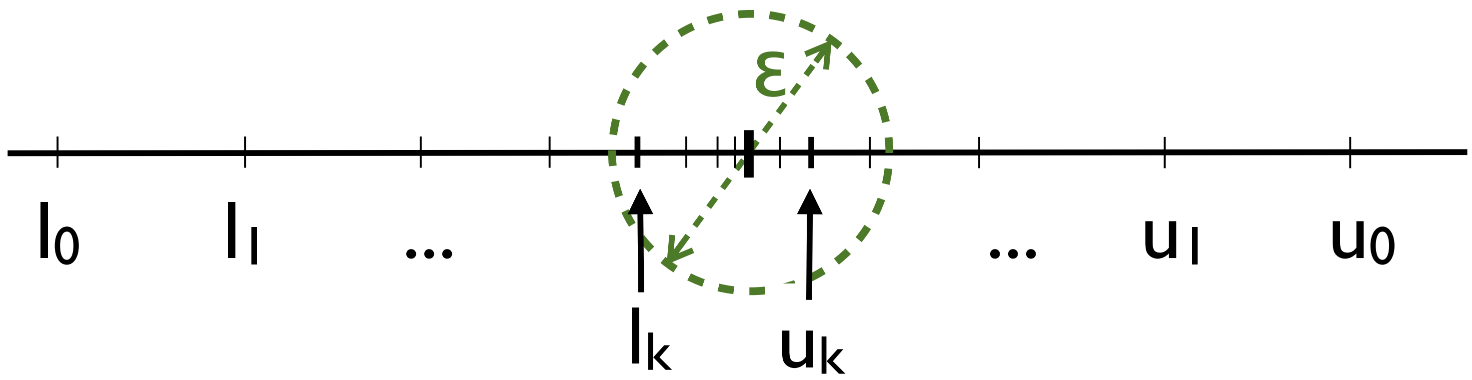
$$u_i = 1 - \sup \{m(X) \mid m \leq_{F_i} \mu \;\&\; m \leq_{F_i} \nu\}$$

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

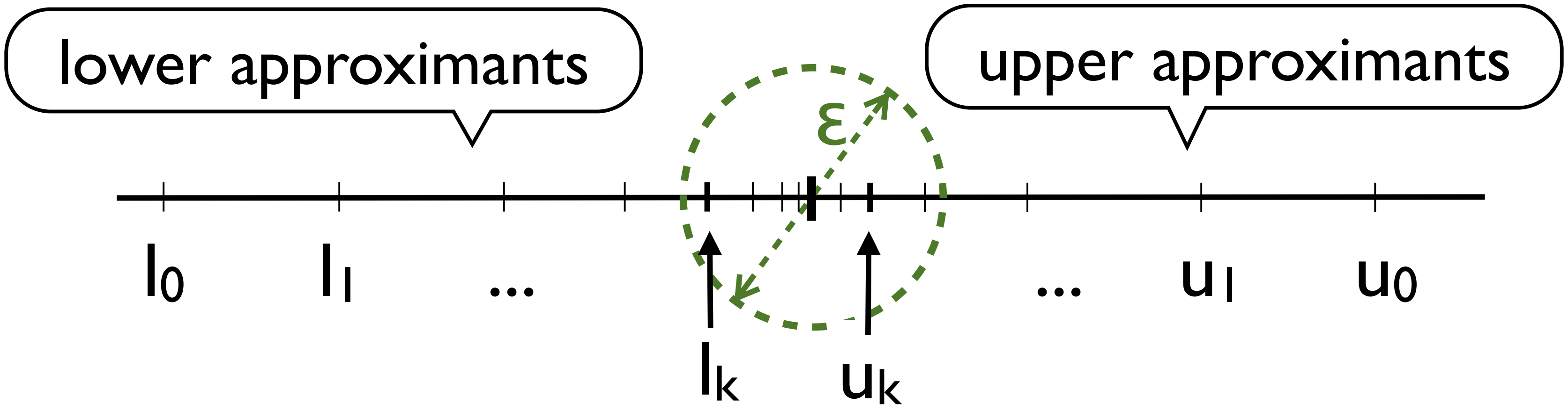
decreasing

limiting

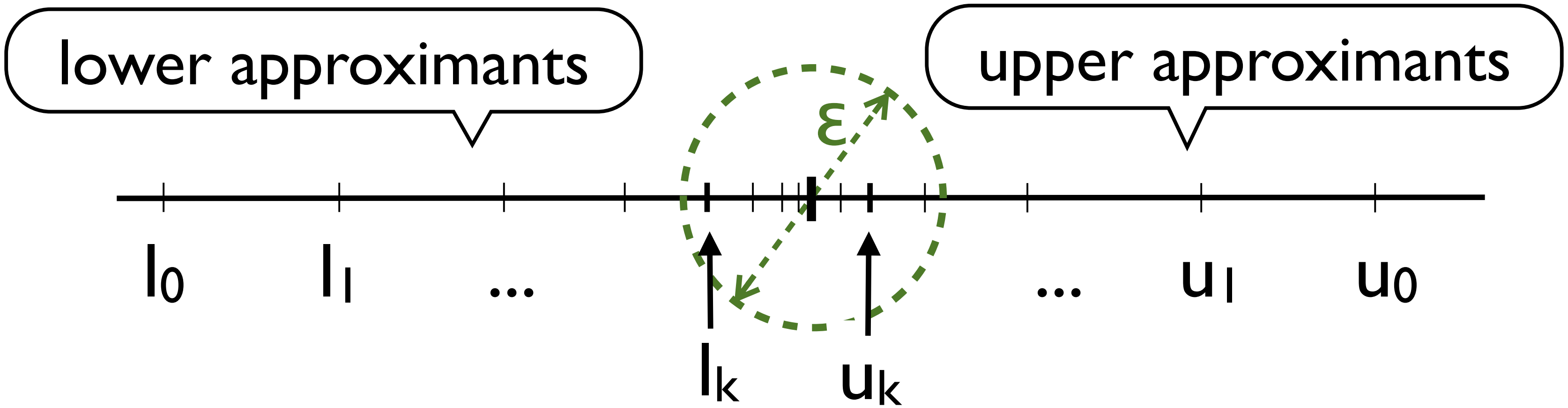
Approximation Algorithm



Approximation Algorithm

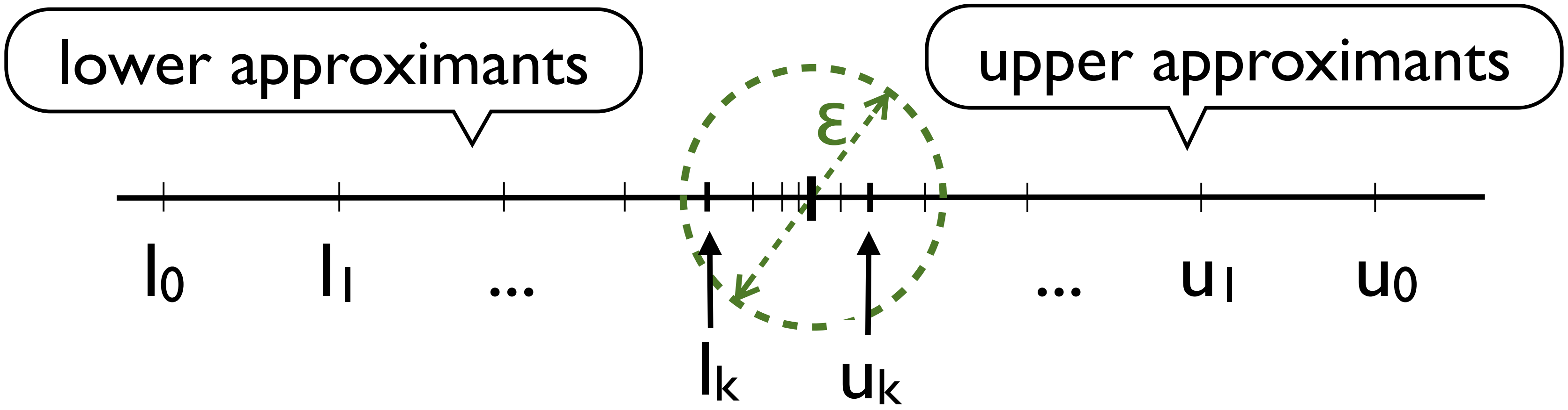


Approximation Algorithm



- Both l_i and u_i are parametric in F_i

Approximation Algorithm



- Both l_i and u_i are parametric in F_i
- If for all $E \in F_i$ $\mu(E)$ and $\nu(E)$ are *computable* then, so are l_i and u_i.

Approximating the Trace Distance

Approximating the Trace Distance

just define the F_i's

Approximating the Trace Distance

just define the F_i's

w.r.t. Trace of Cylinders

$\mathfrak{C}(S_0, [\frac{m_0}{i}, \frac{n_0}{i}], \ldots, [\frac{m_i}{i}, \frac{n_i}{i}], S_{i+1})$ s.t. $m_j < n_j \leq i^2$, $S_j = U_k$ L_k

Approximating the Trace Distance

just define the F_i's

w.r.t. Trace of Cylinders

$$\mathfrak{C}(S_0, [\tfrac{m_0}{i}, \tfrac{n_0}{i}], \ldots, [\tfrac{m_i}{i}, \tfrac{n_i}{i}], S_{i+1}) \quad \text{s.t.} \quad \begin{array}{c} m_j < n_j \leq i^2 \\ S_j = U_k \; L_k \end{array}$$

w.r.t. MTL properties

$$\varphi \coloneqq p \mid \bot \mid \varphi \rightarrow \varphi \mid X^{[\frac{m}{i}, \frac{n}{i}]}\varphi \quad \text{s.t.} \quad \begin{array}{c} m < n \leq i^2 \\ \mathrm{mdepth}(\varphi) \leq i \end{array}$$

Approximating the Trace Distance

just define the F_i's

w.r.t. Trace of Cylinders

$\mathfrak{C}(S_0, [\frac{m_0}{i}, \frac{n_0}{i}], \ldots, [\frac{m_i}{i}, \frac{n_i}{i}], S_{i+1})$ s.t. $m_j < n_j \leq i^2$, $S_j = U_k$ L_k

w.r.t. MTL properties

$\varphi ::= p \mid \bot \mid \varphi \rightarrow \varphi \mid X^{[\frac{m}{i}, \frac{n}{i}]}\varphi$ s.t. $m < n \leq i^2$, $\mathrm{mdepth}(\varphi) \leq i$

w.r.t. Timed Languages

$\mathscr{A} \in$ 1-DTA ...guards $g ::= x \leq \frac{m}{i} \mid x \geq \frac{m}{i} \mid g \wedge g$ $(m \leq i^2)$

Approximating the Trace Distance

just define the F_i's

w.r.t. Trace of Cylinders

$\mathfrak{C}(S_0, [\frac{m_0}{i}, \frac{n_0}{i}], \ldots, [\frac{m_i}{i}, \frac{n_i}{i}], S_{i+1})$ s.t. $m_j < n_j \leq i^2$, $S_j = U_k$ L_k

w.r.t. MTL properties

$\varphi ::= p \mid \bot \mid \varphi \rightarrow \varphi \mid X^{[\frac{m}{i}, \frac{n}{i}]} \varphi$ s.t. $m < n \leq i^2$, $mdepth(\varphi) \leq i$

computable!
Chen et al. [LICS'09]

w.r.t. Timed Languages

$\mathcal{A} \in$ 1-DTA ...guards $g ::= x \leq \frac{m}{i} \mid x \geq \frac{m}{i} \mid g \wedge g \ (m \leq i^2)$

Complexity Results

In terms of the complexity of approximating the trace distance we have the following result

NP-hardness [Lyngsø-Pedersen JCSS'02]

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

Complexity Results

In terms of the complexity of approximating the trace distance we have the following result

NP-hardness [Lyngsø-Pedersen JCSS'02]

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

reduction from the max-clique problem

Concluding Remarks

Concluding Remarks

- A truly genuine connection between trace distance and the model checking problem

Concluding Remarks

- A truly genuine connection between trace distance and the model checking problem
- General results for total variation distance:

Concluding Remarks

- A truly genuine connection between trace distance and the model checking problem
- General results for total variation distance:
  - algebraic representation theorem

Concluding Remarks

- A truly genuine connection between trace distance and the model checking problem
- General results for total variation distance:
  - algebraic representation theorem
  - approximation strategies (& algorithm)

Concluding Remarks

- A truly genuine connection between trace distance and the model checking problem
- General results for total variation distance:
  - algebraic representation theorem
  - approximation strategies (& algorithm)
- A polynomial upper-bound (not shown)

Thank you
for the attention