Converging from Branching to Linear Metrics on MCs

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

29 October 2015 - Cali, Colombia

ICTAC 2015

The focus of the talk

The focus of the talk

- We are interested in ***Quantitative Aspects***

The focus of the talk

- We are interested in ***Quantitative Aspects***
  - ***Models*** - probabilistic, timed, weighted, etc.

The focus of the talk

- We are interested in ***Quantitative Aspects***
  - ***Models*** - probabilistic, timed, weighted, etc.
  - ***Behavior*** - from equivalences to distances

The focus of the talk

- We are interested in ***Quantitative Aspects***
  - ***Models*** - probabilistic, timed, weighted, etc.
  - ***Behavior*** - from equivalences to distances
  - ***Formal Verification*** - quantitative Model Checking

The focus of the talk

- We are interested in ***Quantitative Aspects***
  - ***Models*** - probabilistic, timed, weighted, etc.
  - ***Behavior*** - from equivalences to distances
  - ***Formal Verification*** - quantitative Model Checking
- in particular: ***Linear-time Properties***

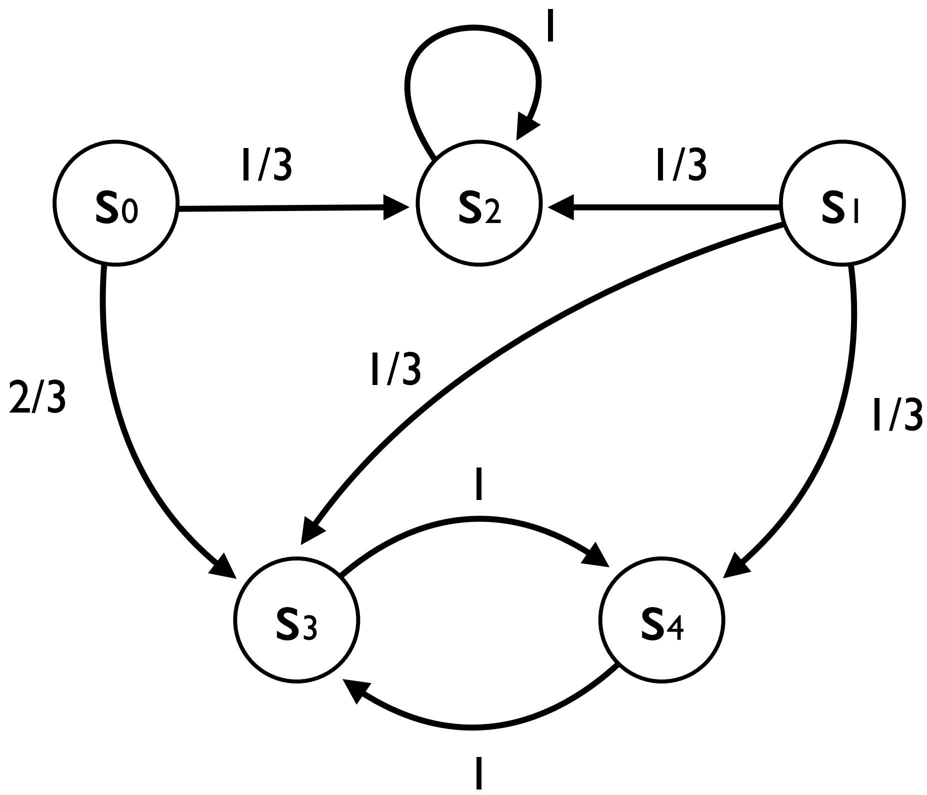
The focus of the talk

- We are interested in ***Quantitative Aspects***
  - ***Models*** - probabilistic, timed, weighted, etc.
  - ***Behavior*** - from equivalences to distances
  - ***Formal Verification*** - quantitative Model Checking
- in particular: ***Linear-time Properties***
  - observables are execution runs (no internal access!)

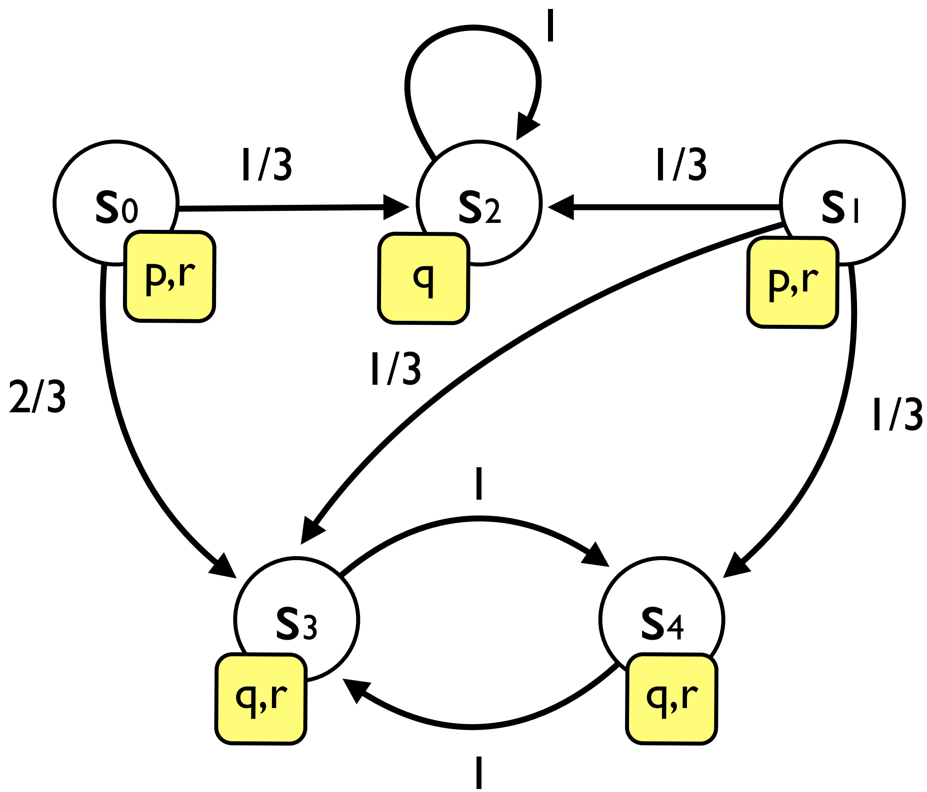
The focus of the talk

- We are interested in ***Quantitative Aspects***
    - ***Models*** - probabilistic, timed, weighted, etc.
    - ***Behavior*** - from equivalences to distances
    - ***Formal Verification*** - quantitative Model Checking
- in particular: ***Linear-time Properties***
    - observables are execution runs (no internal access!)
    - **Why?** --systems biology, machine learning, artificial intelligence, security, etc.

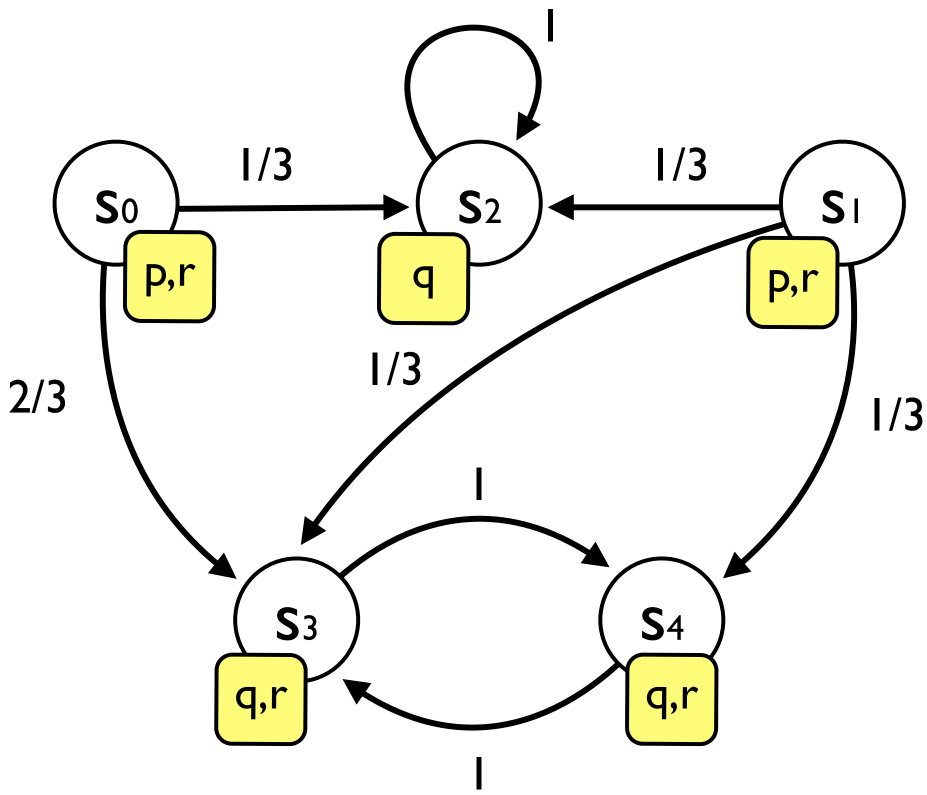
Markov Chains



Markov Chains



Markov Chains

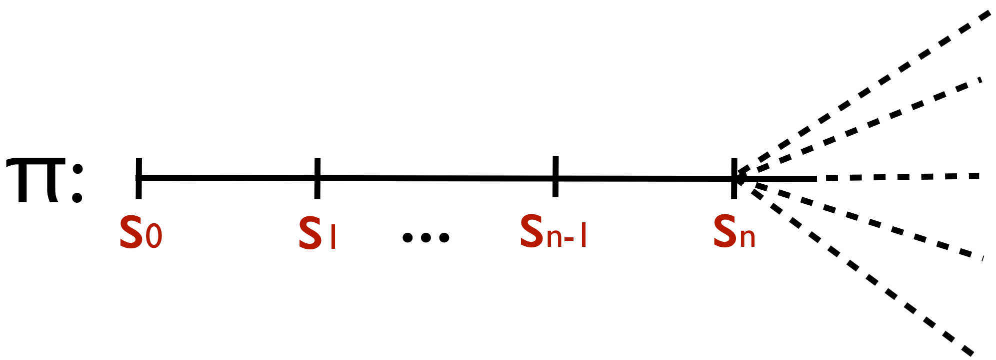


Given an initial state, MCs can be interpreted as “machines” that emit infinite traces of states with a certain probability

Measurable Events

Measurable Events

Cylinder set (or cone)
(with prefix $s_0 \ldots s_n \in S_0 \ldots S_n$)

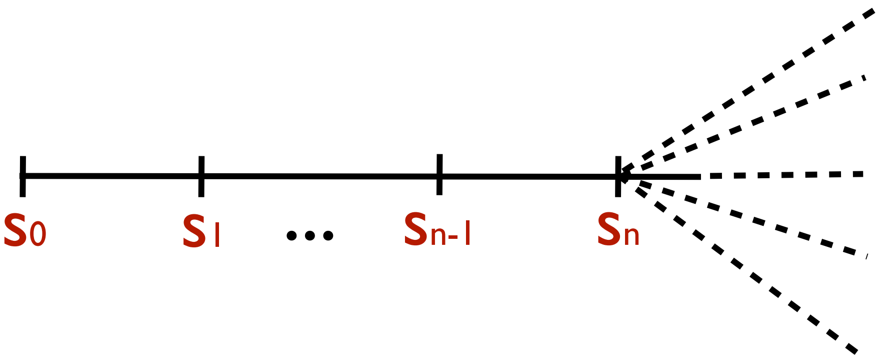


$\in \mathfrak{C}(S_0 \ldots S_n)$

Measurable Events

Cylinder set (or cone)
(with prefix $s_0 \ldots s_n \in S_0 \ldots S_n$)

π:



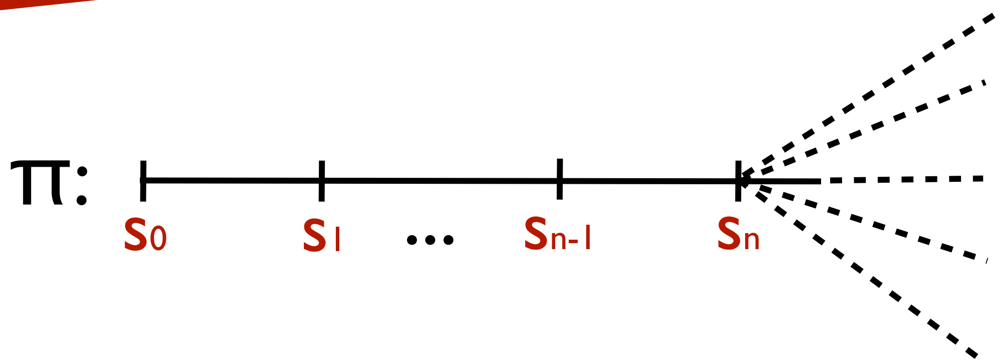
$\in \mathfrak{C}(S_0 \ldots S_n)$

$P(s)(\mathfrak{C}(S_0 \ldots S_n)) =$ "probability that, *starting from s*, the MC emits a path with prefix in $S_0 \ldots S_n$"

Measurable Events

Cylinder-set construction

Cylinder set (or cone)
(with prefix $s_0 \ldots s_n \in S_0 \ldots S_n$)



$\in \mathfrak{C}(S_0 \ldots S_n)$

$P(s)(\mathfrak{C}(S_0 \ldots S_n)) =$ "probability that, *starting from s*, the MC emits a path with prefix in $S_0 \ldots S_n$"

(Pnueli)

Linear Temporal Logic

Atomic prop. — Next — Until

$$\varphi \coloneqq p \mid \bot \mid \varphi \rightarrow \varphi \mid X\varphi \mid \varphi U \varphi$$

(Pnueli)

Linear Temporal Logic

Atomic prop. — Next — Until

$$\varphi \coloneqq p \mid \bot \mid \varphi \rightarrow \varphi \mid X\varphi \mid \varphi U \varphi$$

Semantics of a formula

$$[\varphi] = \{\pi \mid \pi \models \varphi\}$$

(Pnueli)

Linear Temporal Logic

Atomic prop. — Next — Until

$$\varphi := p \mid \bot \mid \varphi \rightarrow \varphi \mid X\varphi \mid \varphi U \varphi$$

Semantics of a formula

$$[\varphi] = \{\pi \mid \pi \models \varphi\}$$

with usual satisfiability relation

(Pnueli)

Linear Temporal Logic

Atomic prop. — Next — Until

$$\varphi := p \mid \bot \mid \varphi \to \varphi \mid X\varphi \mid \varphi U \varphi$$

Semantics of a formula

measurable event!

$$[\varphi] = \{\pi \mid \pi \models \varphi\}$$

with usual satisfiability relation

Probabilistic Model Checking

Probabilistic Model Checking

On probabilistic systems we cannot verify strong assertions such as "*the system will never fail*"...

Probabilistic Model Checking

On probabilistic systems we cannot verify strong assertions such as "*the system will never fail*"...

$$P(s)([\varphi]) = ?$$

What is the probabability that the MC with initial state s satisfies the formula φ?

Approximate verification

Approximate verification

- Model Checking does not scale to large systems (even after model reduction, symbolic tecn., partial-order reduction, etc.)

Approximate verification

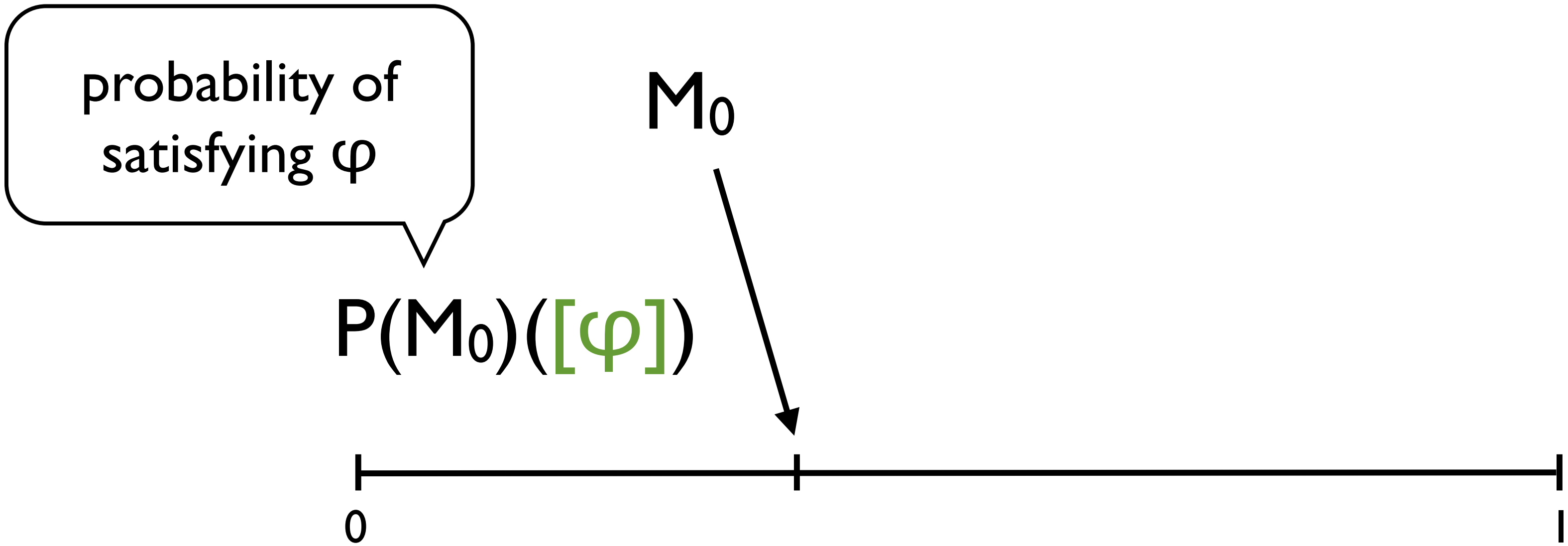
- Model Checking does not scale to large systems (even after model reduction, symbolic tecn., partial-order reduction, etc.)
- One should reduce the accuracy of the model, ...hence introduce an error

Approximate verification

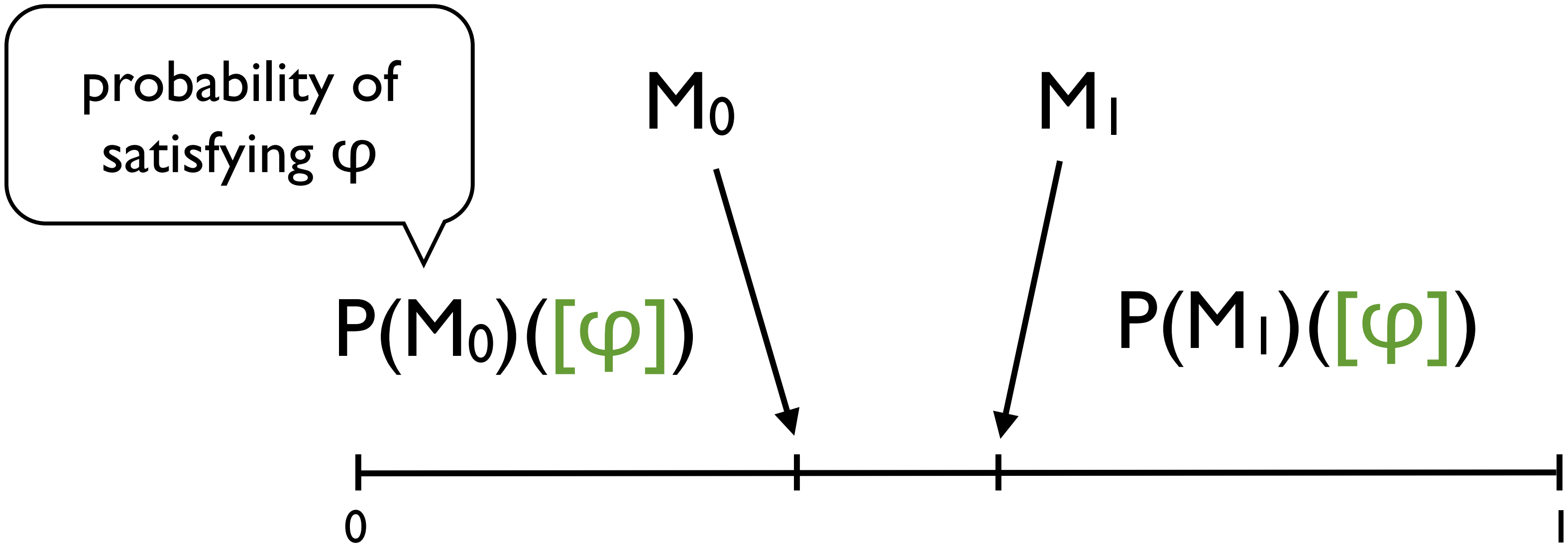
- Model Checking does not scale to large systems (even after model reduction, symbolic tecn., partial-order reduction, etc.)
- One should reduce the accuracy of the model, ...hence introduce an error
- **Proposed solution:**
  Behavioral metrics to quatify the error

A distance for approx. Model Checking

A distance for approx. Model Checking

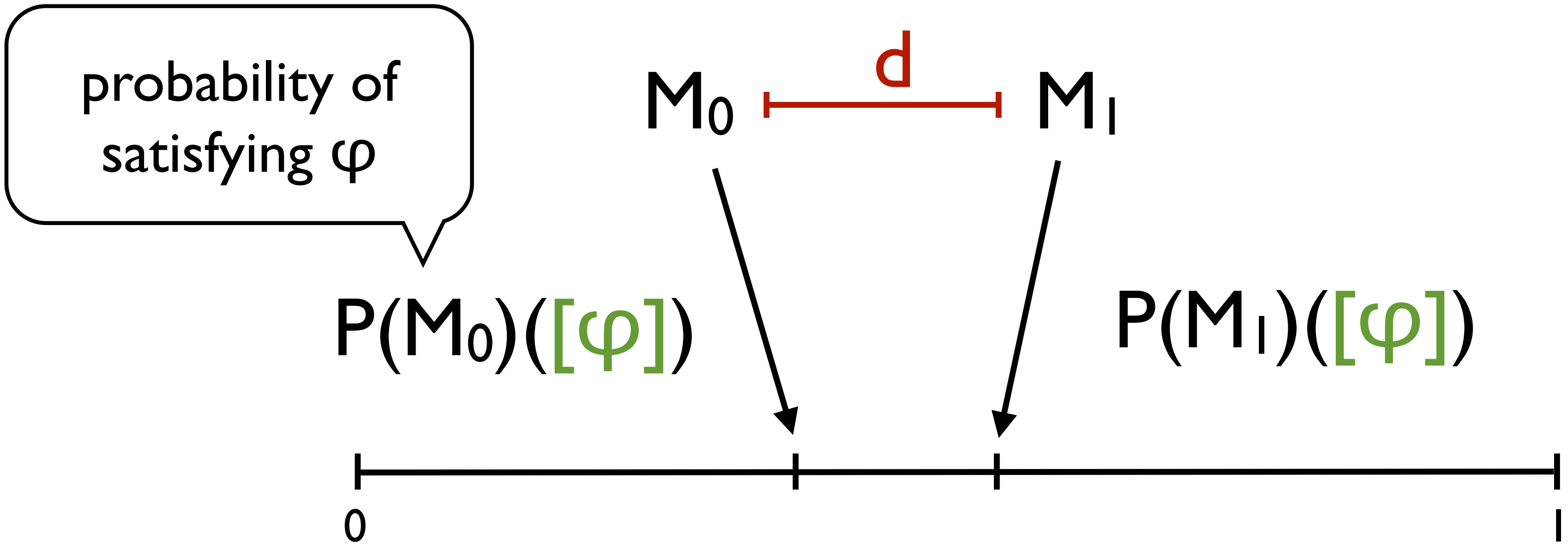


A distance for approx. Model Checking



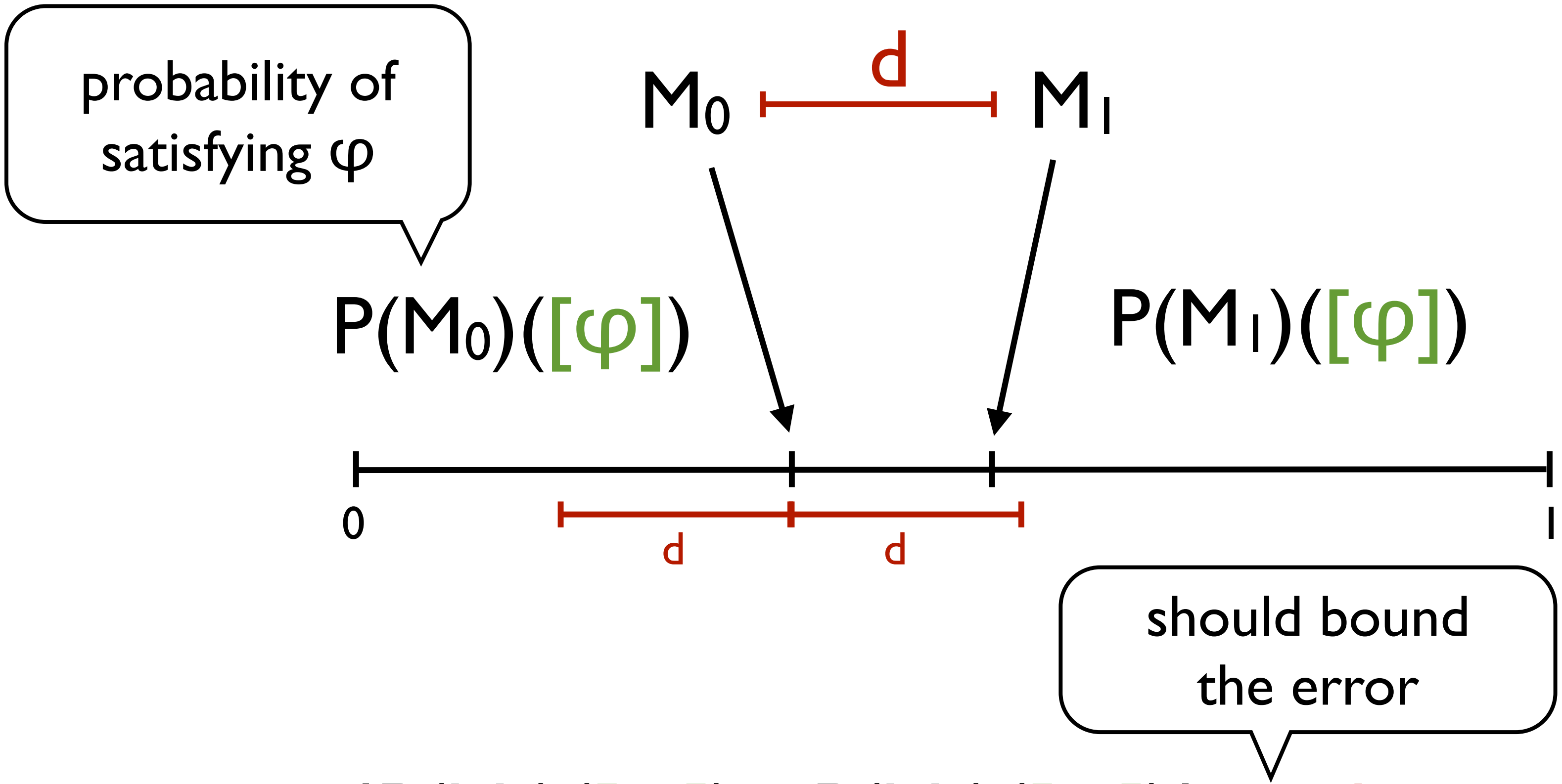
$$|P(M_0)([\varphi]) - P(M_1)([\varphi])|$$

A distance for approx. Model Checking



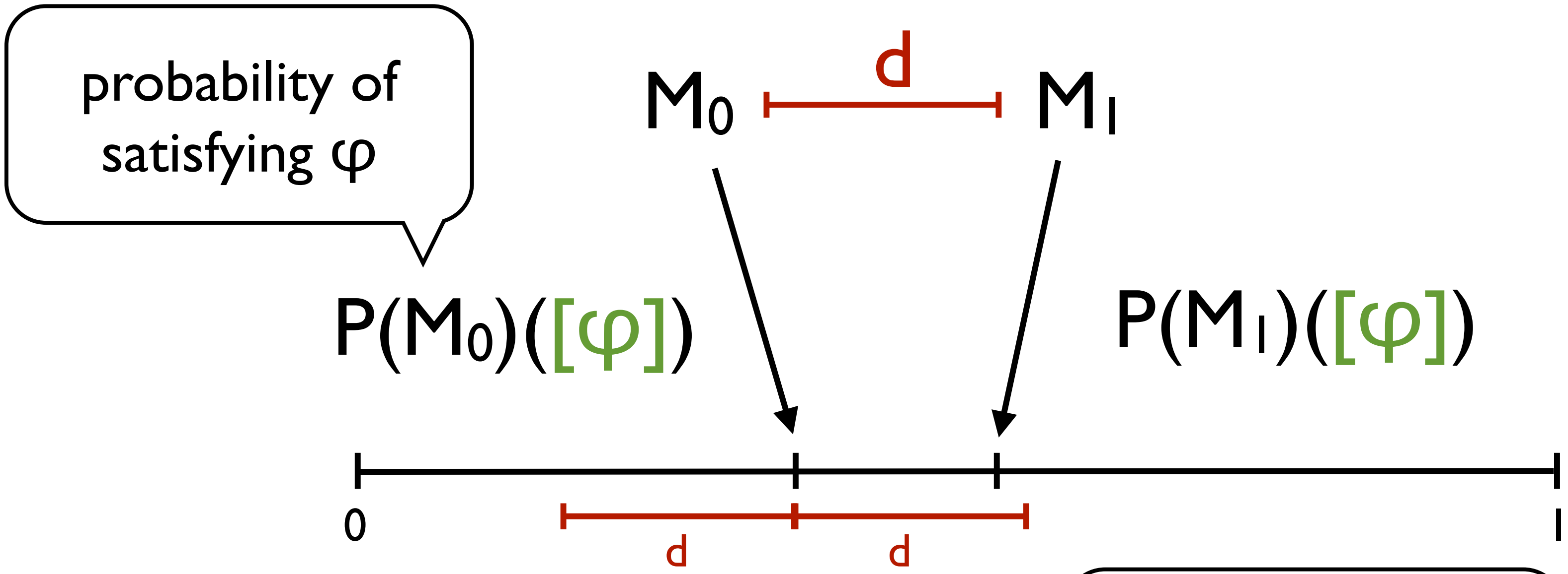
$$|P(M_0)([\varphi]) - P(M_1)([\varphi])|$$

A distance for approx. Model Checking



$$|P(M_0)([\varphi]) - P(M_1)([\varphi])| \leq d$$

A distance for approx. Model Checking



should bound the error

$$|P(M_0)([\varphi]) - P(M_1)([\varphi])| \leq d$$

for all formulas!

Two logical distances

Two logical distances

the LTL distance

$$LTL(s,t) = \sup_{\varphi \in LTL} |P(s)([\varphi]) - P(t)([\varphi])|$$

Two logical distances

the LTL distance

$$LTL(s,t) = \sup_{\varphi \in LTL} |P(s)([\varphi]) - P(t)([\varphi])|$$

the LTL^{-x} distance

$$LTL^{-x}(s,t) = \sup_{\varphi \in LTL^{-x}} |P(s)([\varphi]) - P(t)([\varphi])|$$

Two logical distances

the LTL distance

$$LTL(s,t) = \sup_{\varphi \in LTL} |P(s)([\varphi]) - P(t)([\varphi])|$$

the LTL^{-x} distance

$$LTL^{-x}(s,t) = \sup_{\varphi \in LTL^{-x}} |P(s)([\varphi]) - P(t)([\varphi])|$$

LTL without next operator

Two logical distances

Three natural questions

Q1: Can we compute the two metrics?

Q2: Can we compute them exactly?
If not, can we approximate them
to any arbitrary precision?

Q3: What about complexity?

Characterizations

Trace distance

$$T(s,t) = \sup_{E \in \sigma(\mathcal{T})} |P(s)(E) - P(t)(E)|$$

Stutter-trace distance

$$ST(s,t) = \sup_{E \in \sigma(\mathcal{ST})} |P(s)(E) - P(t)(E)|$$

Characterizations

Trace distance

$$T(s,t) = \sup_{E \in \sigma(\mathcal{T})} |P(s)(E) - P(t)(E)|$$

Events up-to trace equivalence

$$ST(s,t) = \sup_{E \in \sigma(\mathcal{ST})} |P(s)(E) - P(t)(E)|$$

Characterizations

Trace distance

$$T(s,t) = \sup_{E \in \sigma(\mathcal{T})} |P(s)(E) - P(t)(E)|$$

Stutter-trace distance

$$ST(s,t) = \sup_{E \in \sigma(\mathcal{ST})} |P(s)(E) - P(t)(E)|$$

Events up-to stutter trace equivalence

Characterizations

Trace distance

$$T(s,t) = \sup_{E \in \sigma(\mathcal{T})} |P(s)(E) - P(t)(E)|$$

Stutter-trace distance

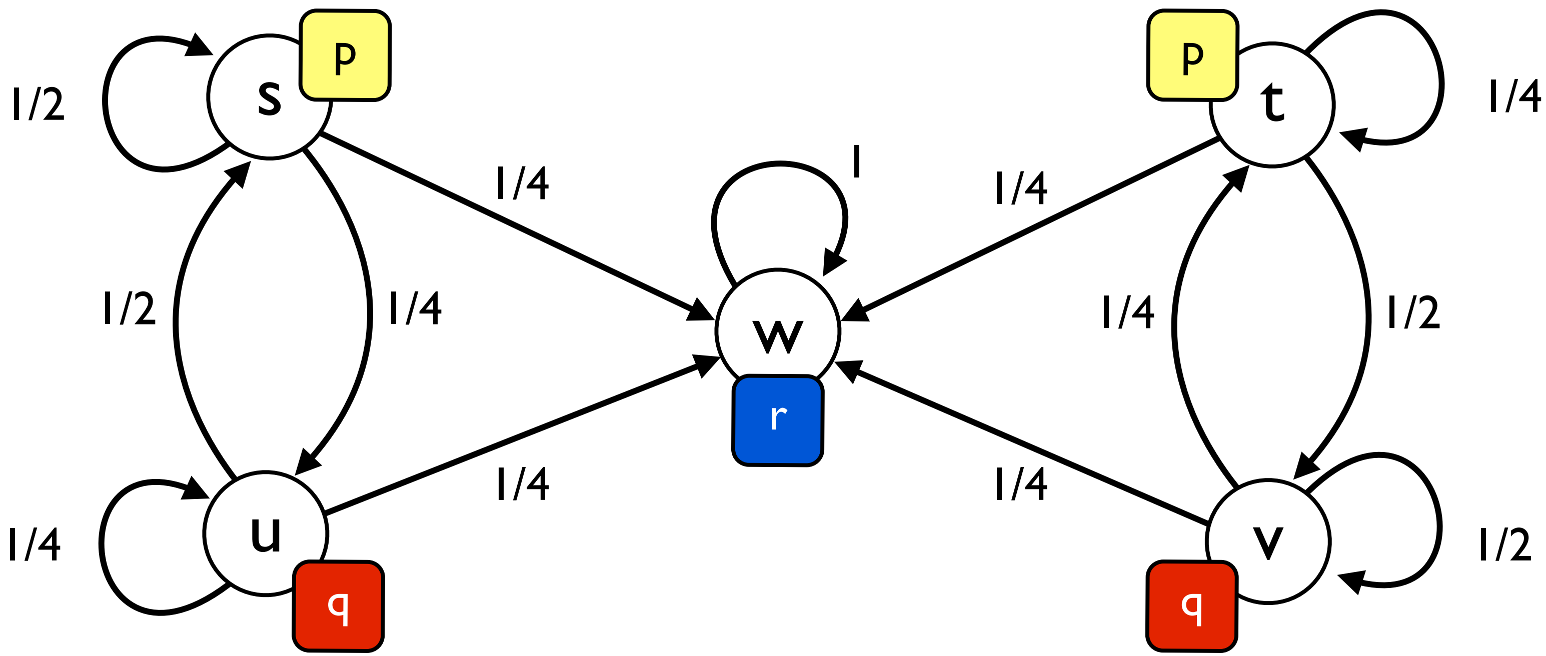
$$ST(s,t) = \sup_{E \in \sigma(\mathcal{ST})} |P(s)(E) - P(t)(E)|$$

Characterization Theorem

$LTL(s,t) = T(s,t)$ and $LTL^{-x}(s,t) = ST(s,t)$

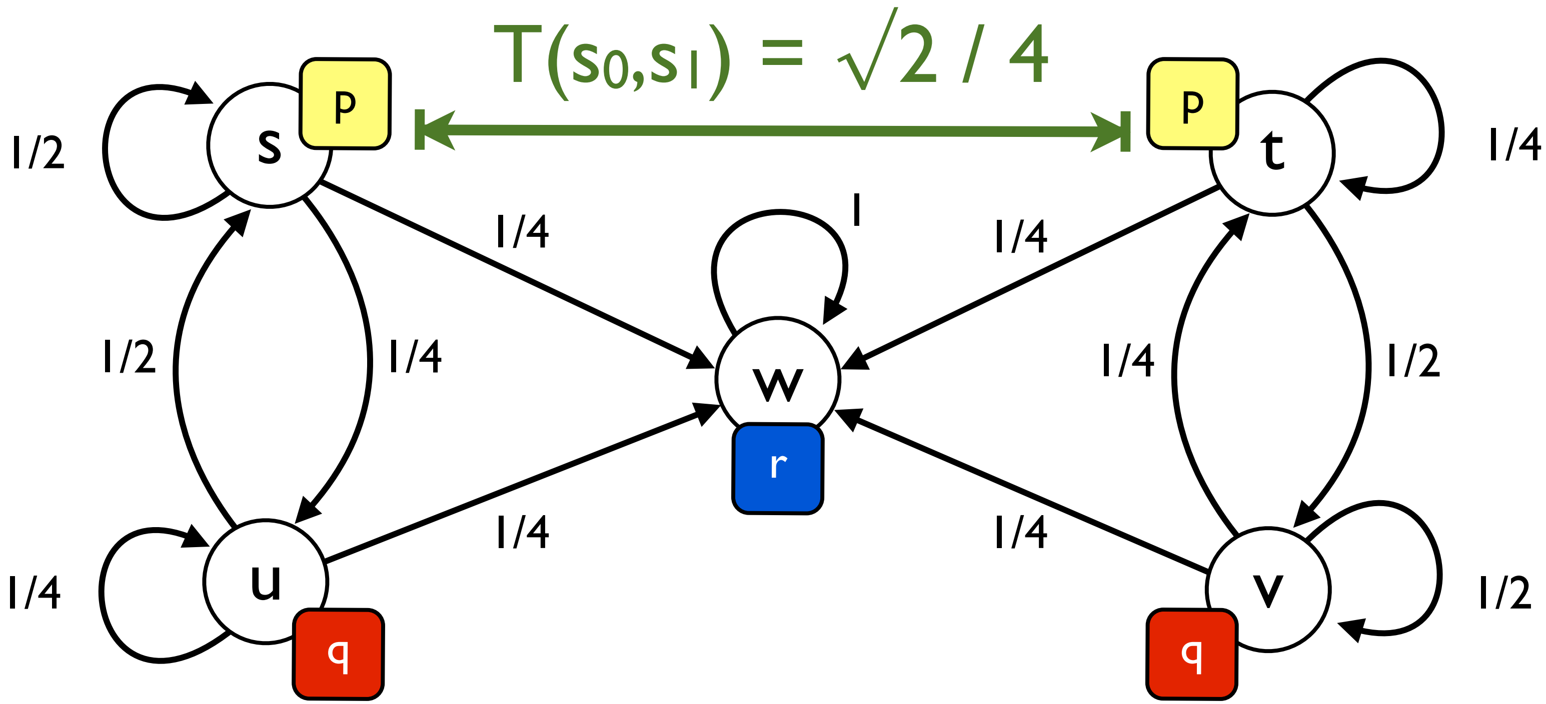
A tiny yet tricky example

(from Chen-Kiefer LICS'14)



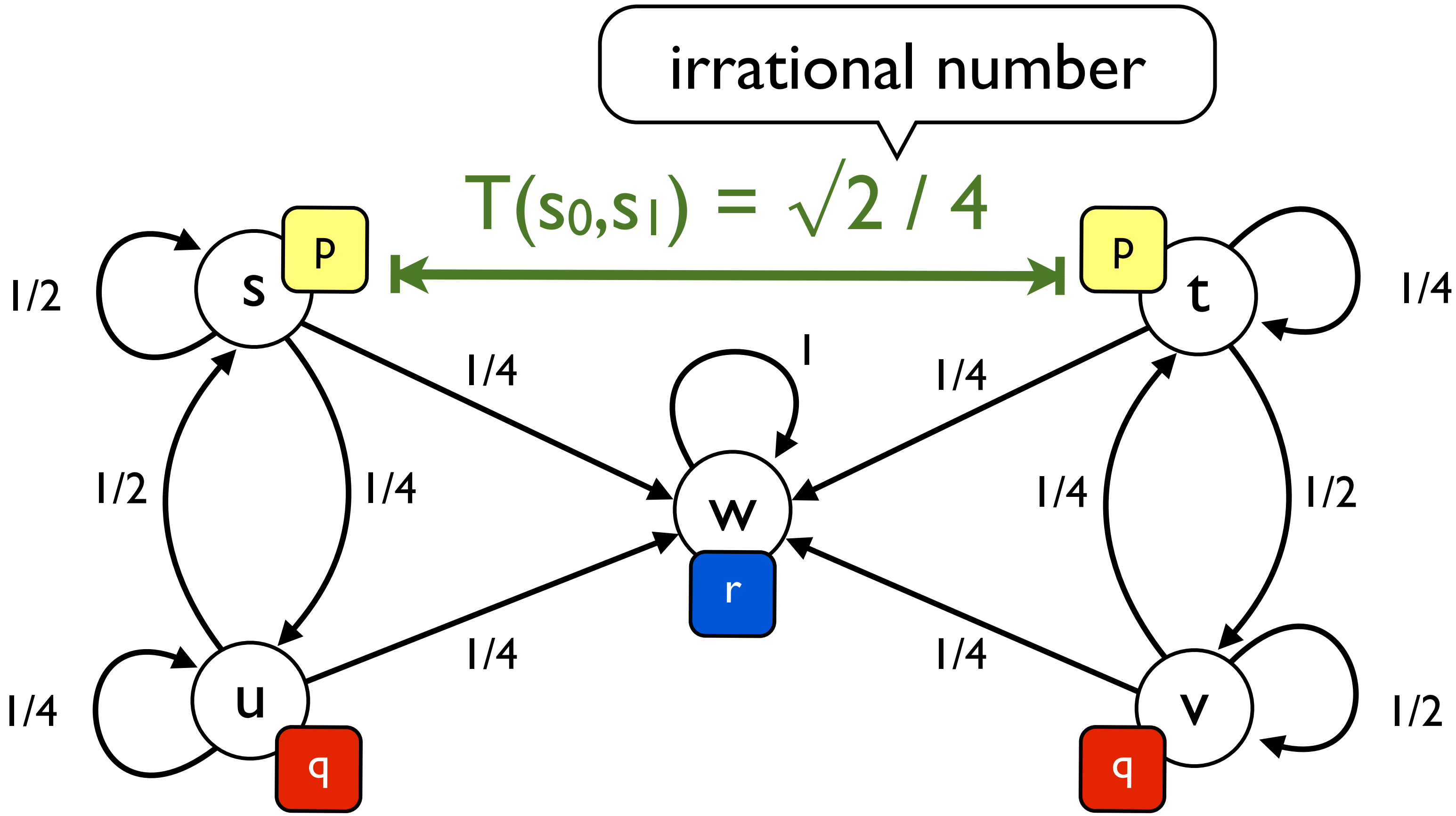
A tiny yet tricky example

(from Chen-Kiefer LICS'14)



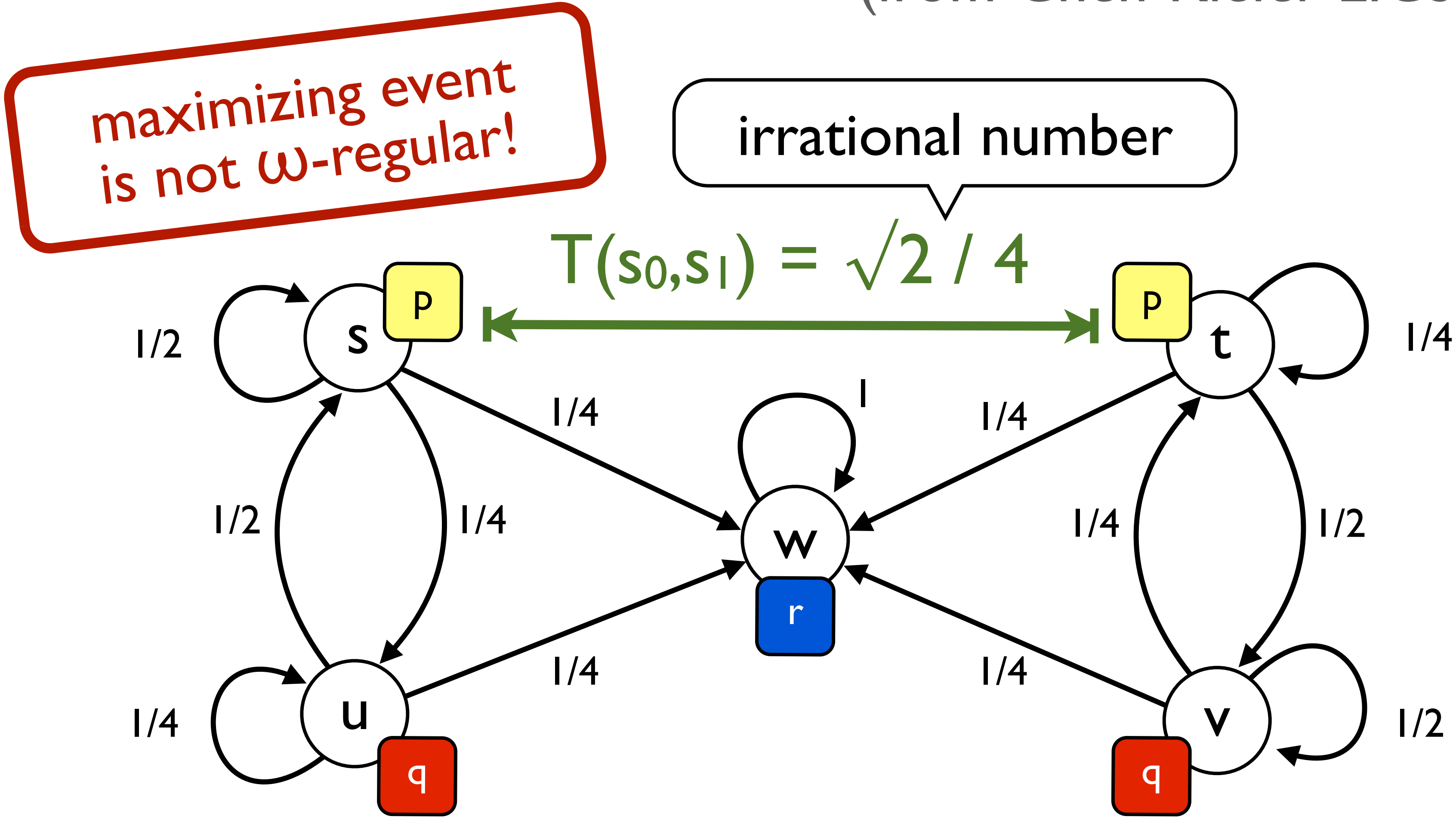
A tiny yet tricky example

(from Chen-Kiefer LICS'14)



A tiny yet tricky example

(from Chen-Kiefer LICS'14)



Direct Consequences

Direct Consequences

- There is no maximizing formula

Direct Consequences

- There is no maximizing formula
- Decidability is still an open problem

Direct Consequences

- There is no maximizing formula
- Decidability is still an open problem
- The threshold problem is NP-hard
(i.e., whether the distance exceeds a given threshold - Lyngsø-Pedersen JCSS'02)

Direct Consequences

- There is no maximizing formula
- Decidability is still an open problem
- The threshold problem is NP-hard (i.e., whether the distance exceeds a given threshold - Lyngsø-Pedersen JCSS'02)

Q: Can we approximate the logical/trace distances up to any arbitrary precision?

Approximation Algorithm

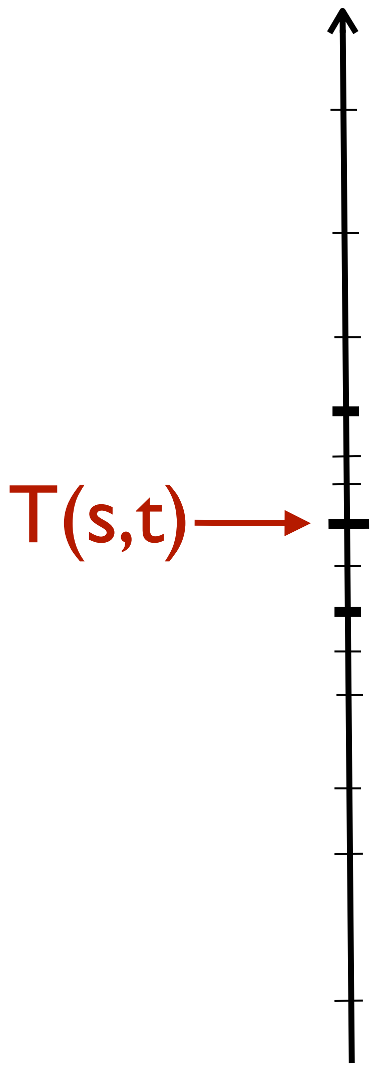
(in the slides only for the Trace Distance)

generalizes / improves
Chen-Kiefer LICS'14

Approximation Algorithm

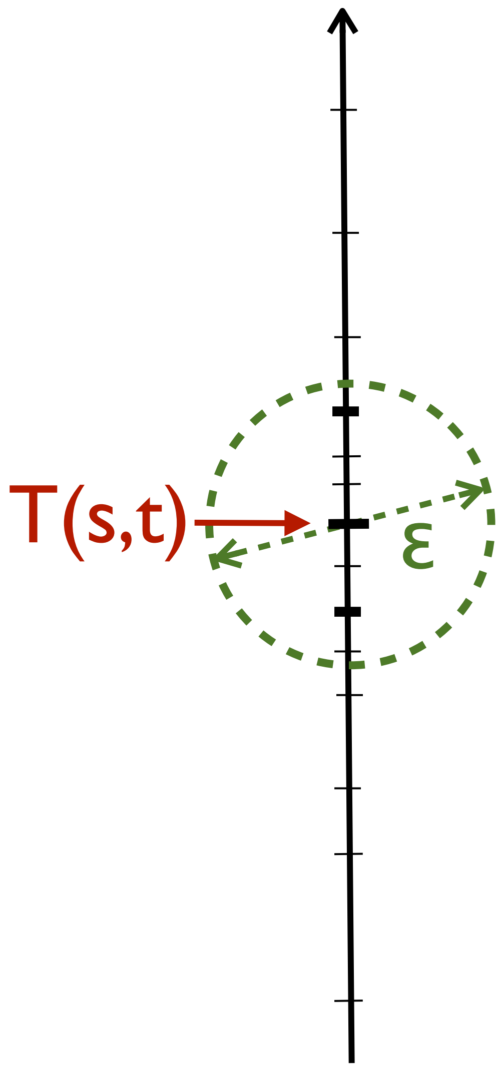
(in the slides only for the Trace Distance)

Approximation Schema
(general idea)

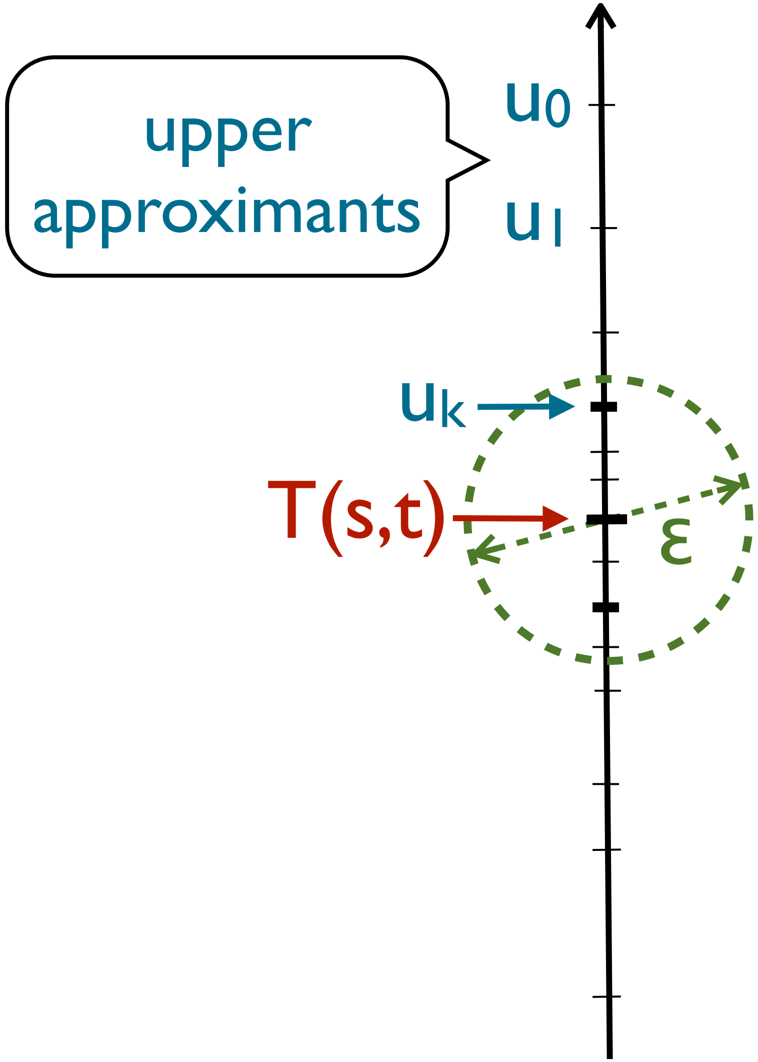


Approximation Schema

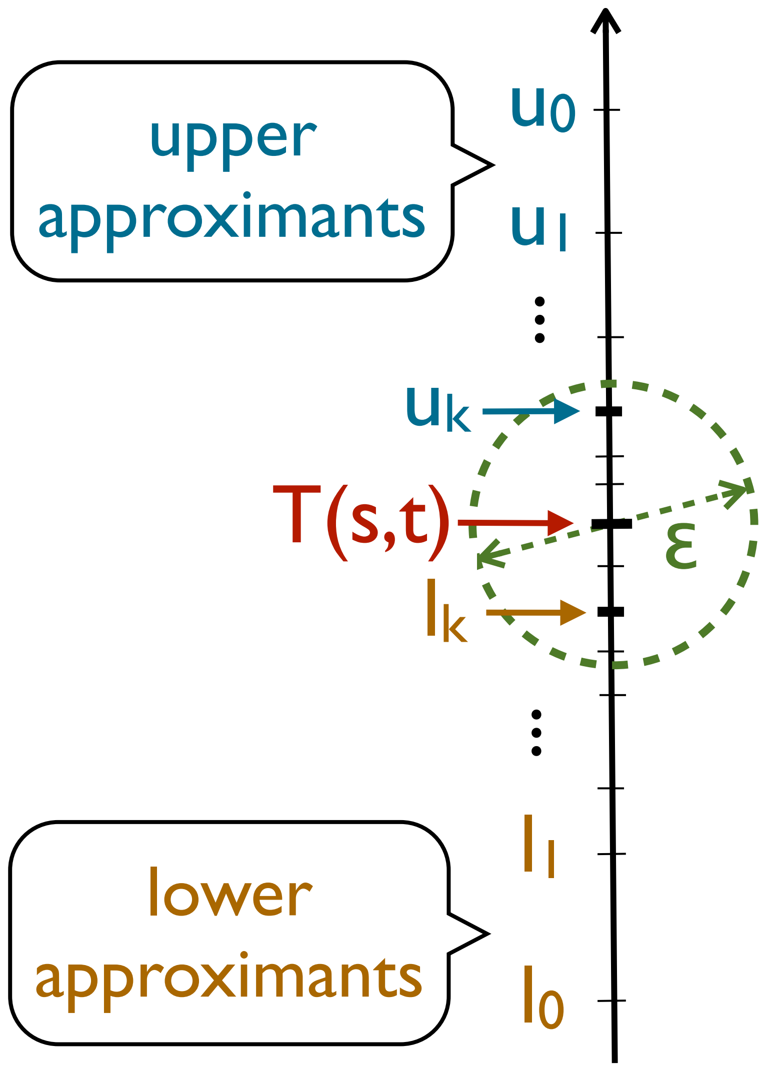
(general idea)



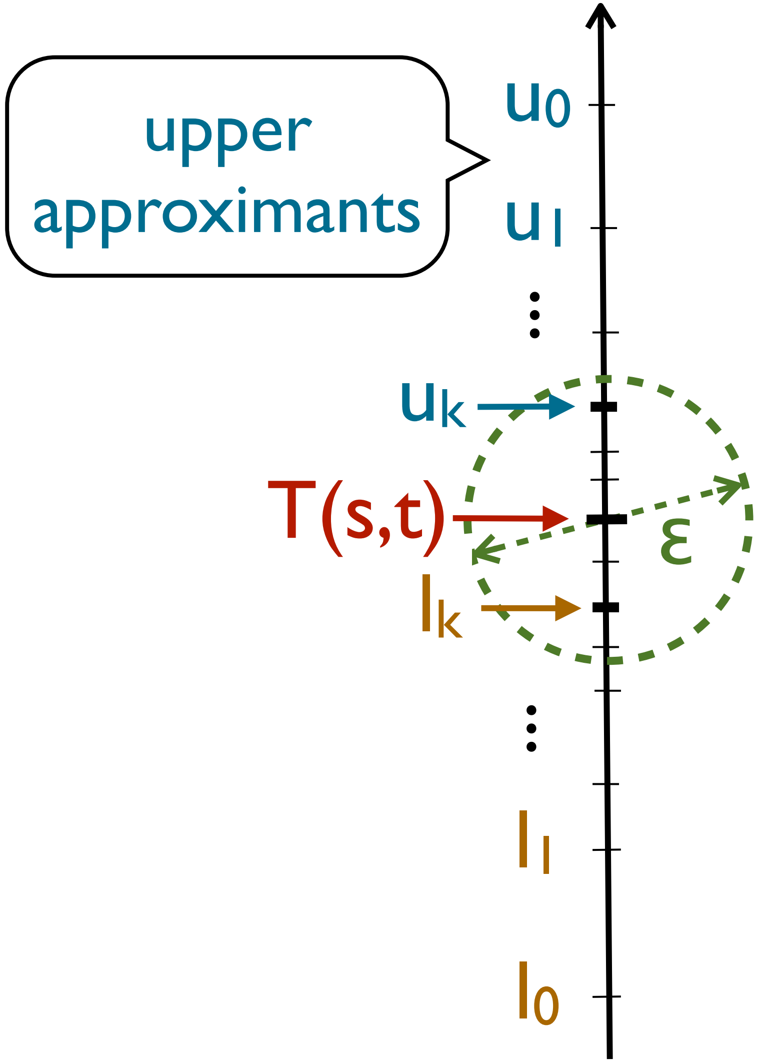
Approximation Schema (general idea)



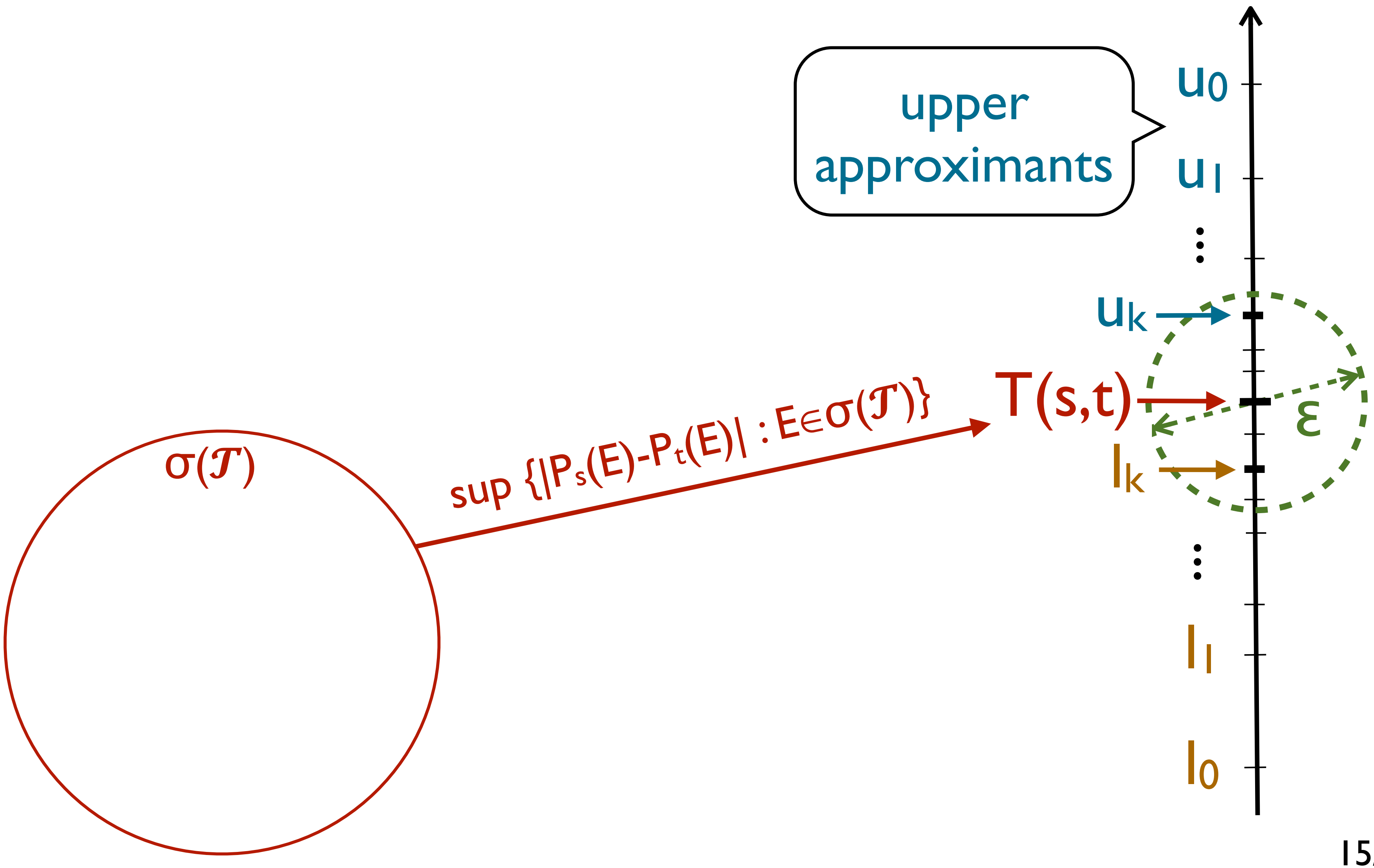
Approximation Schema
(general idea)



Approximation Schema (general idea)

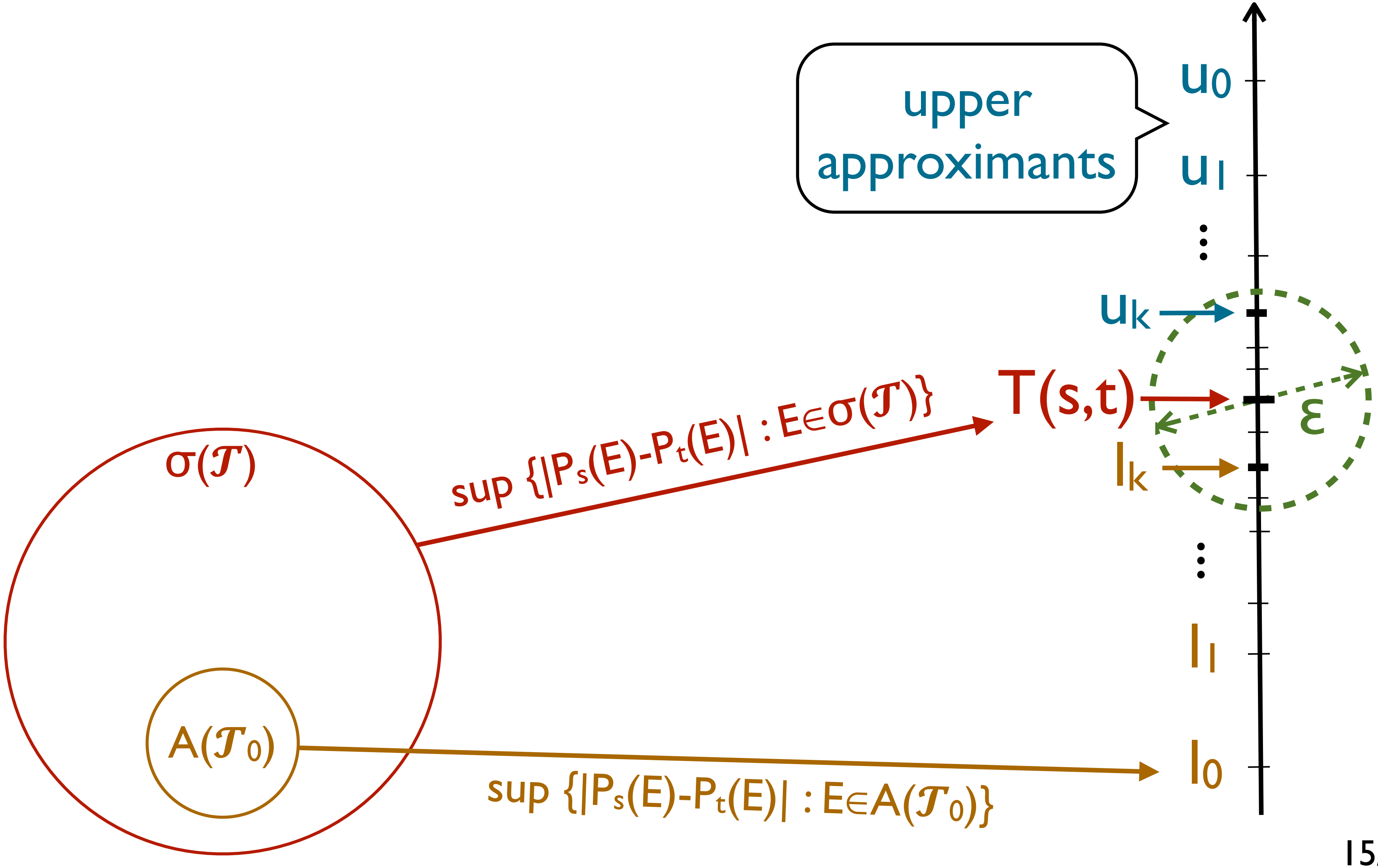


Approximation Schema (general idea)

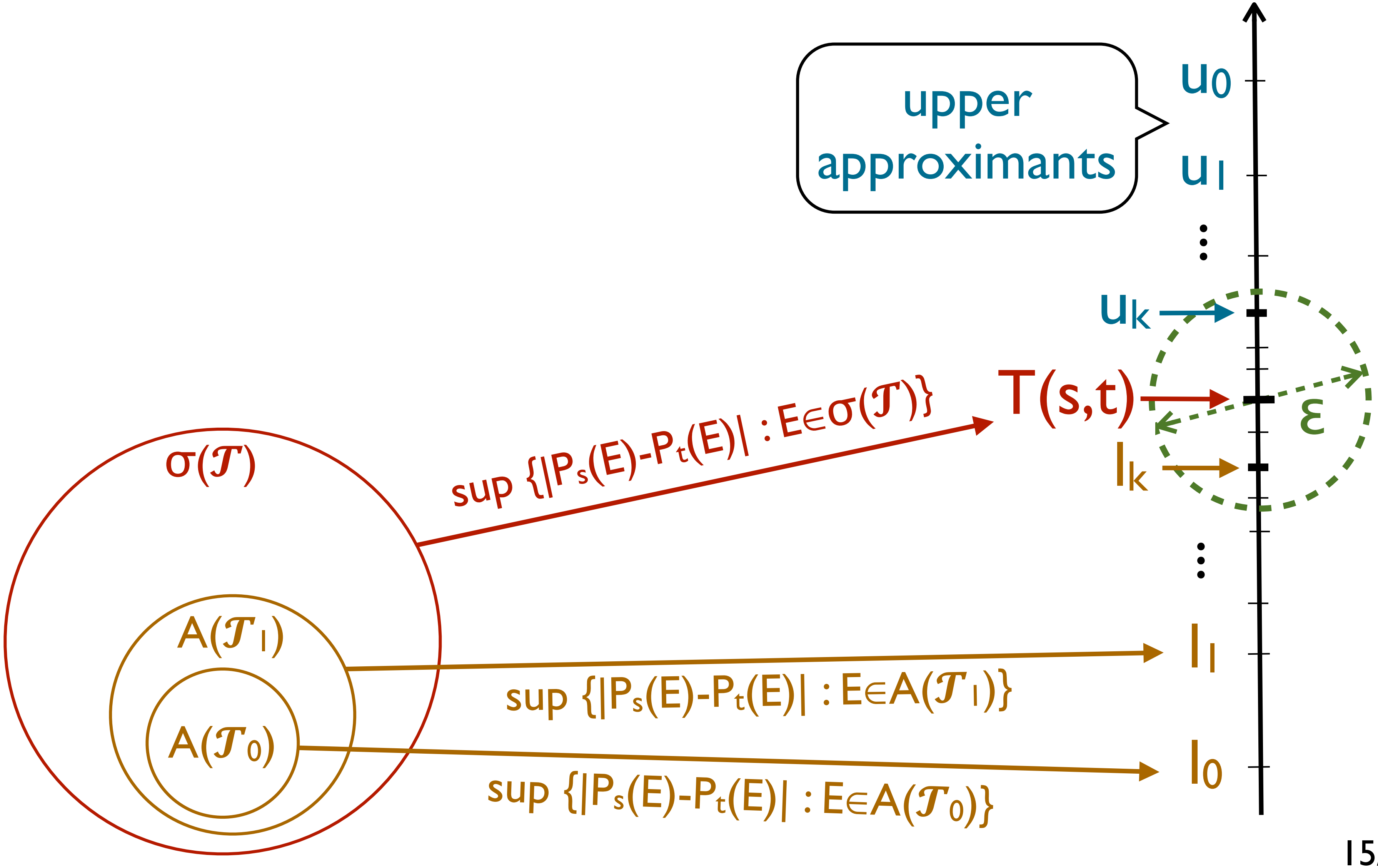


(general idea)

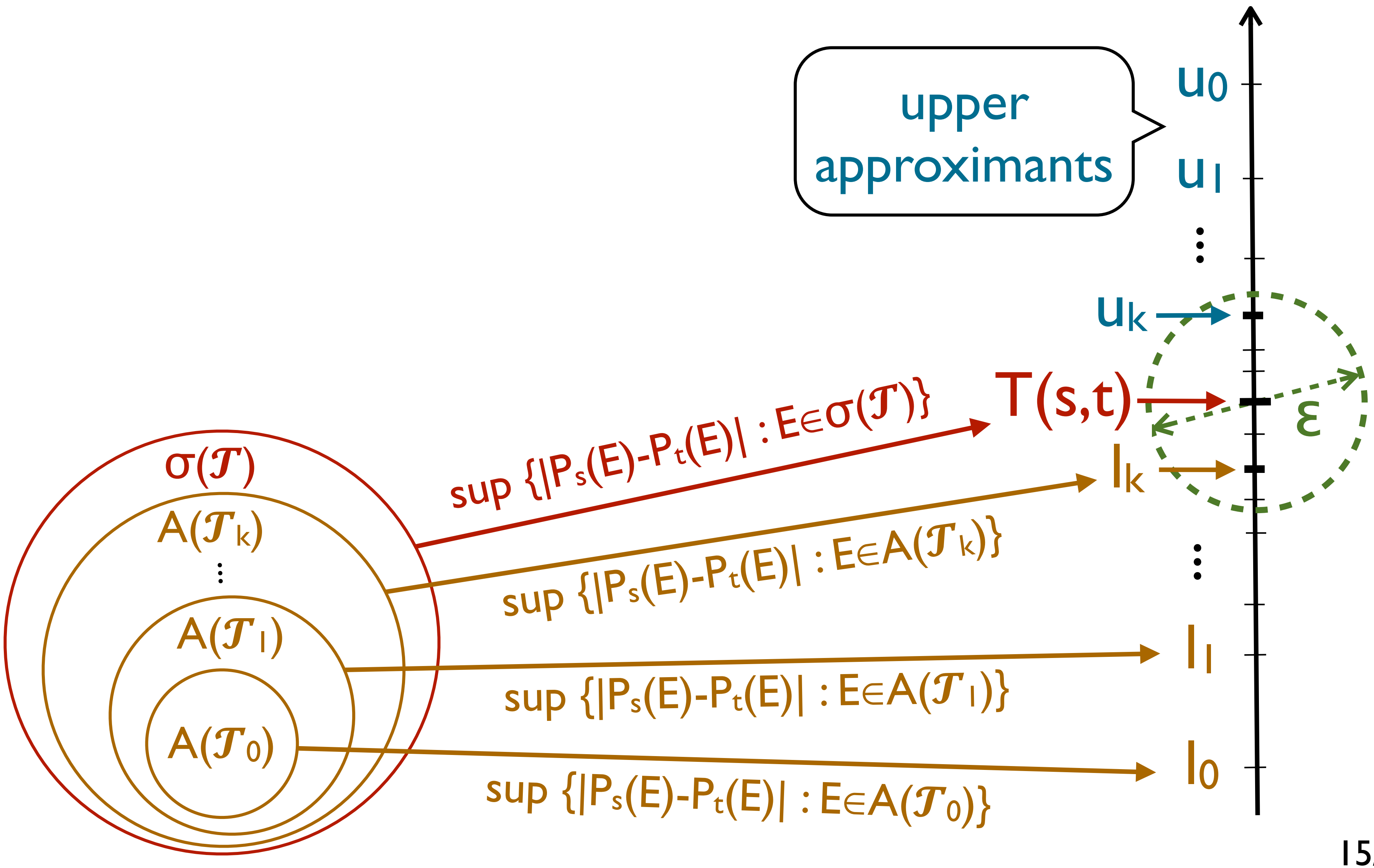
Approximation Schema



Approximation Schema (general idea)



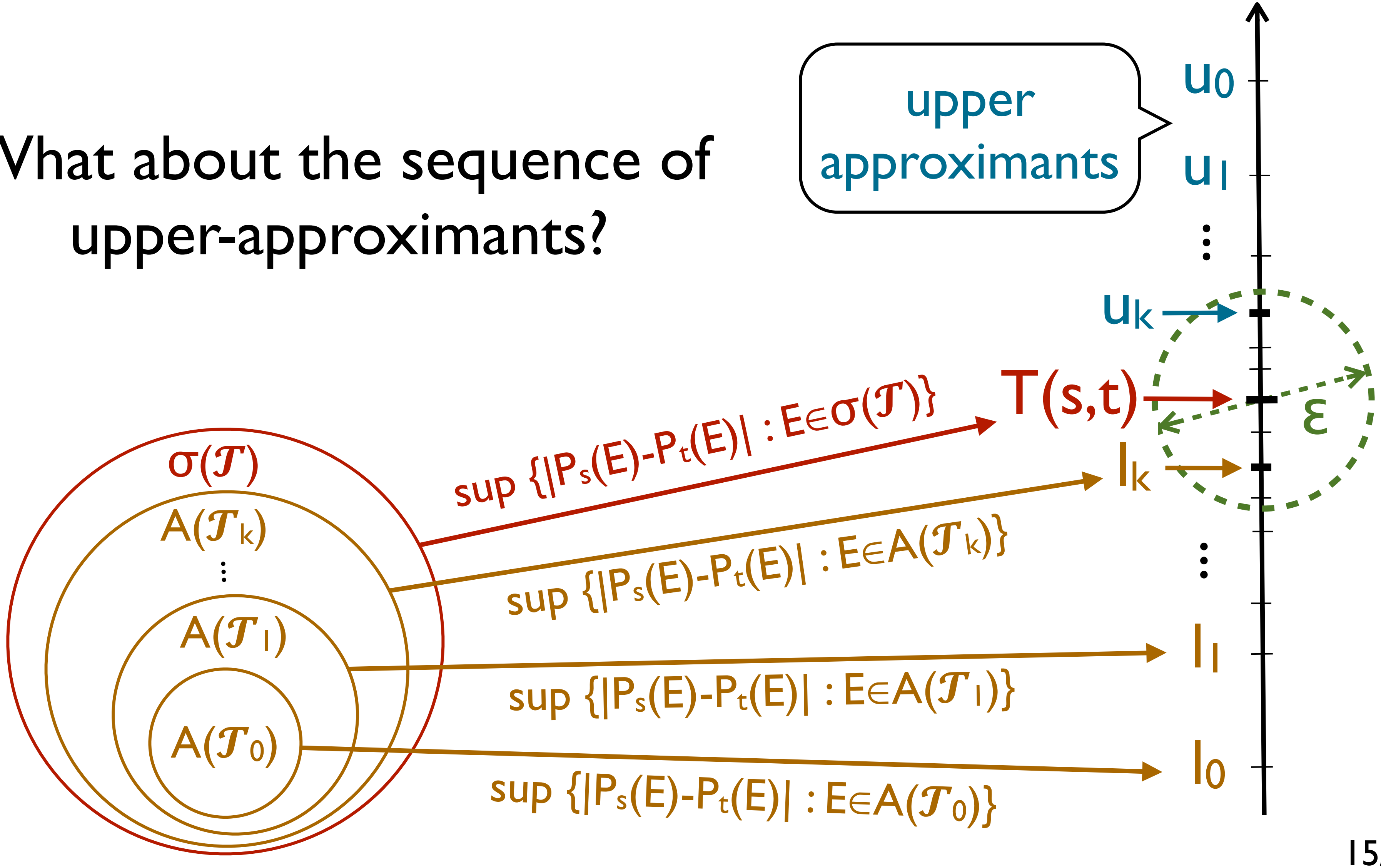
Approximation Schema
(general idea)



(general idea)

Approximation Schema

What about the sequence of upper-approximants?



Coupling Characterization

(as total variation distance)

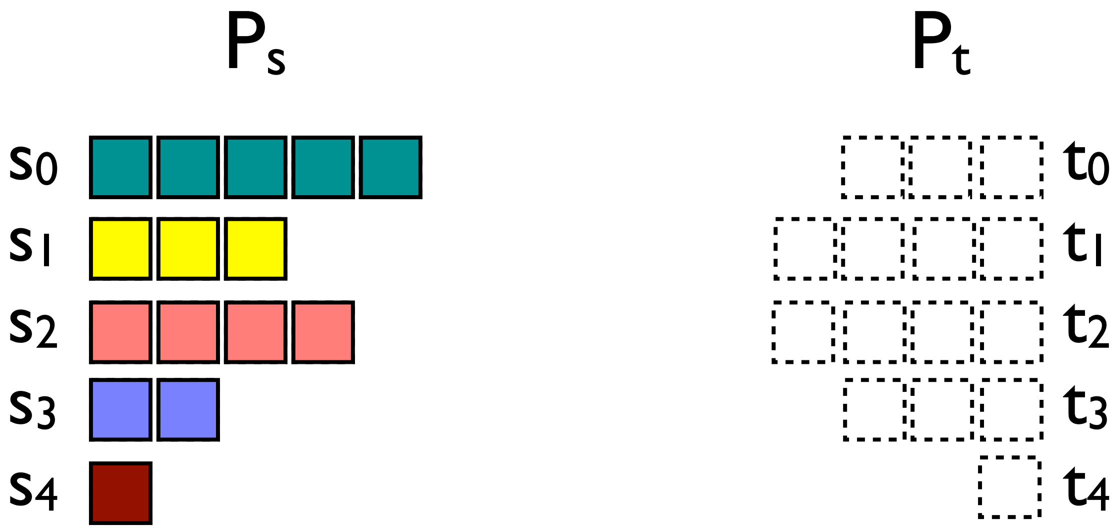
$$T(s,t) = \min \{w(\neq) \mid w \in \Omega(P(s),P(t))\}$$

Coupling Characterization

(as total variation distance)

$$T(s,t) = \min \{w(\neq) \mid w \in \Omega(P(s),P(t))\}$$

Coupling as a transportation schedule...

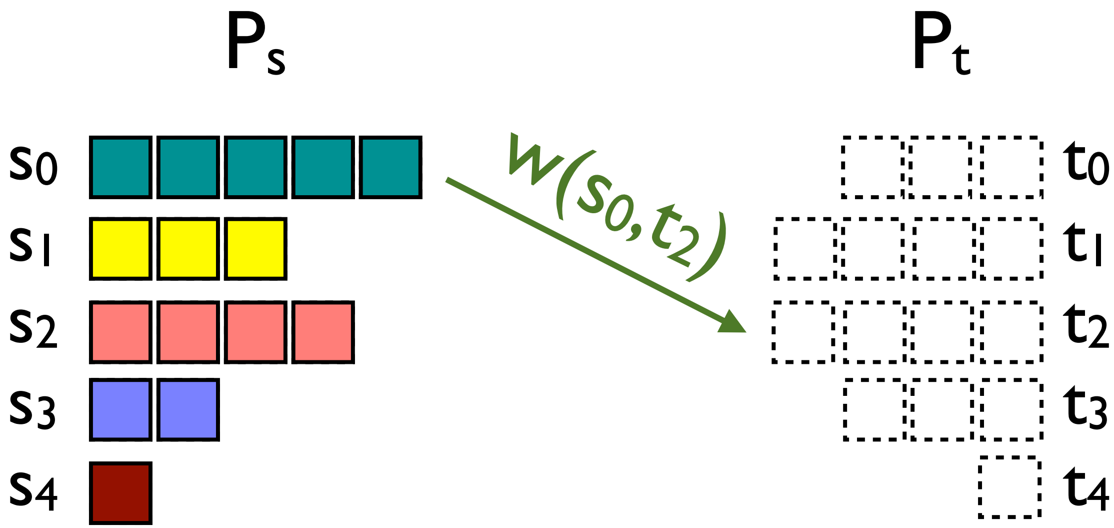


Coupling Characterization

(as total variation distance)

$$T(s,t) = \min \{w(\neq) \mid w \in \Omega(P(s),P(t))\}$$

Coupling as a transportation schedule...

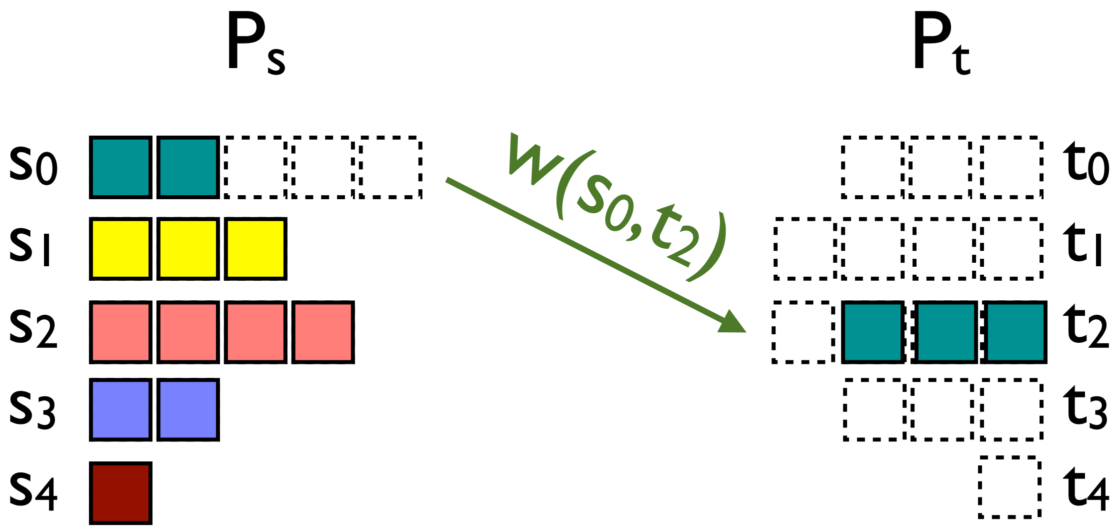


Coupling Characterization

(as total variation distance)

$$T(s,t) = \min \{w(\neq) \mid w \in \Omega(P(s),P(t))\}$$

Coupling as a transportation schedule...



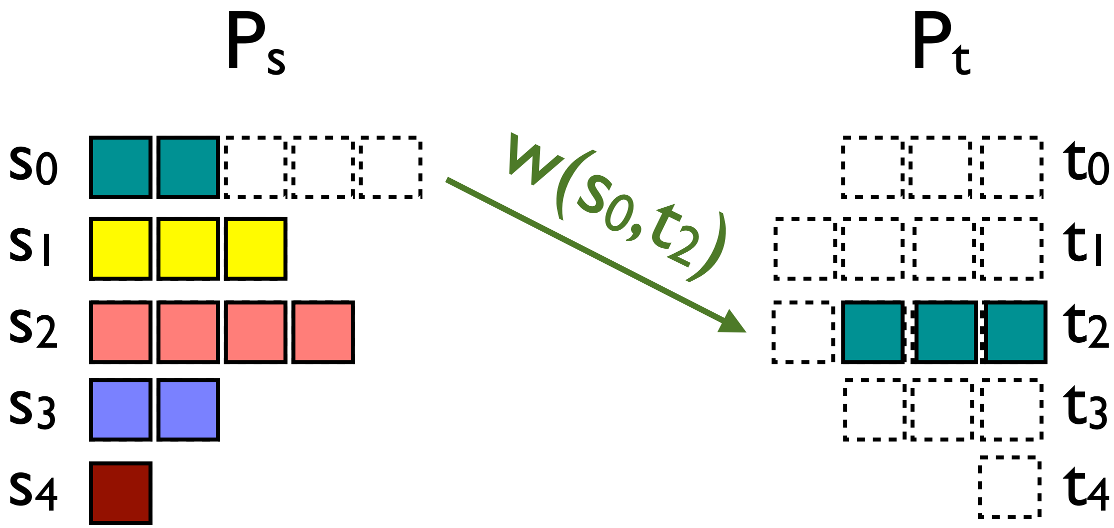
Coupling Characterization

(as total variation distance)

trace inequivalence

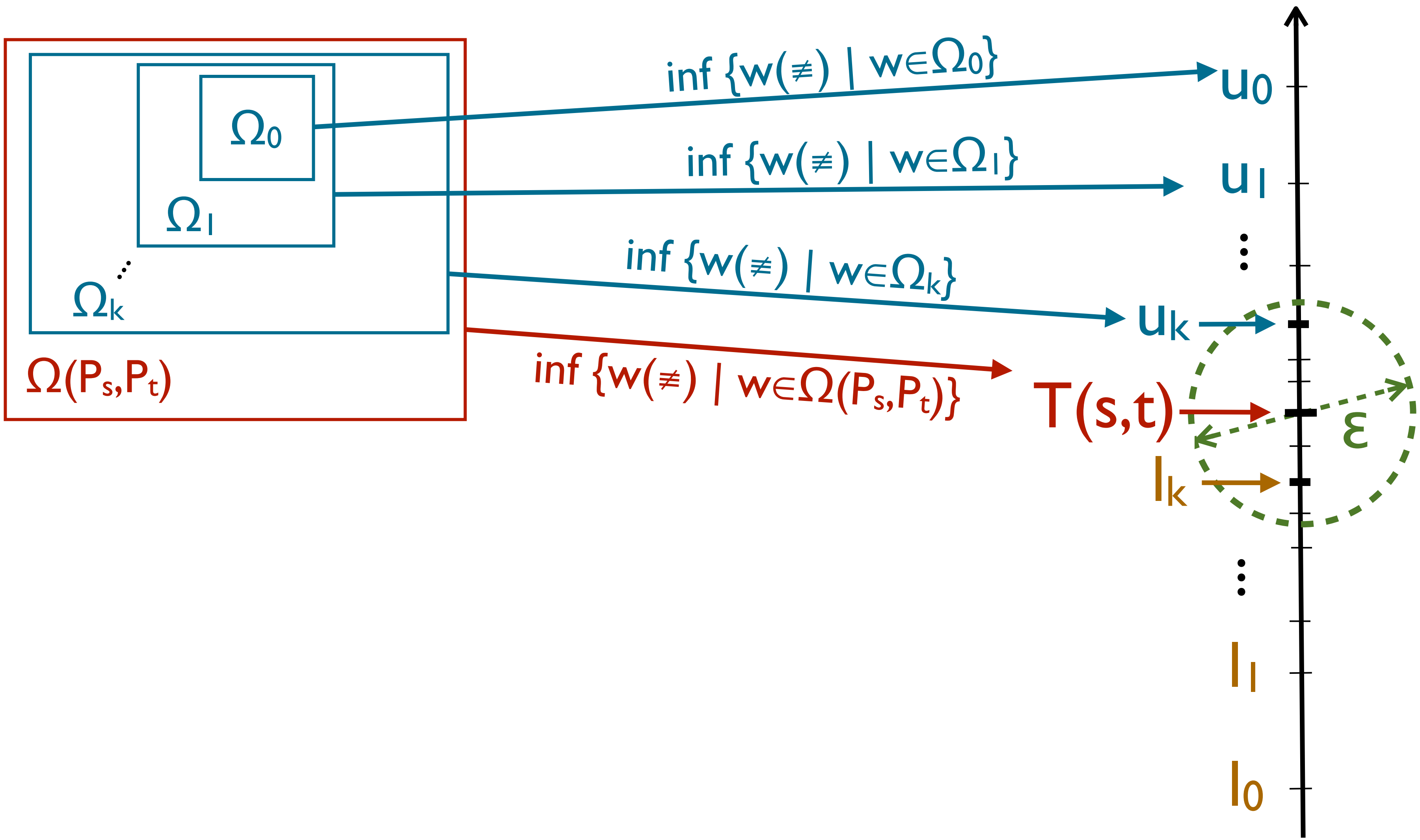
$$T(s,t) = \min \{ w(\neq) \mid w \in \Omega(P(s),P(t)) \}$$

Coupling as a transportation schedule...



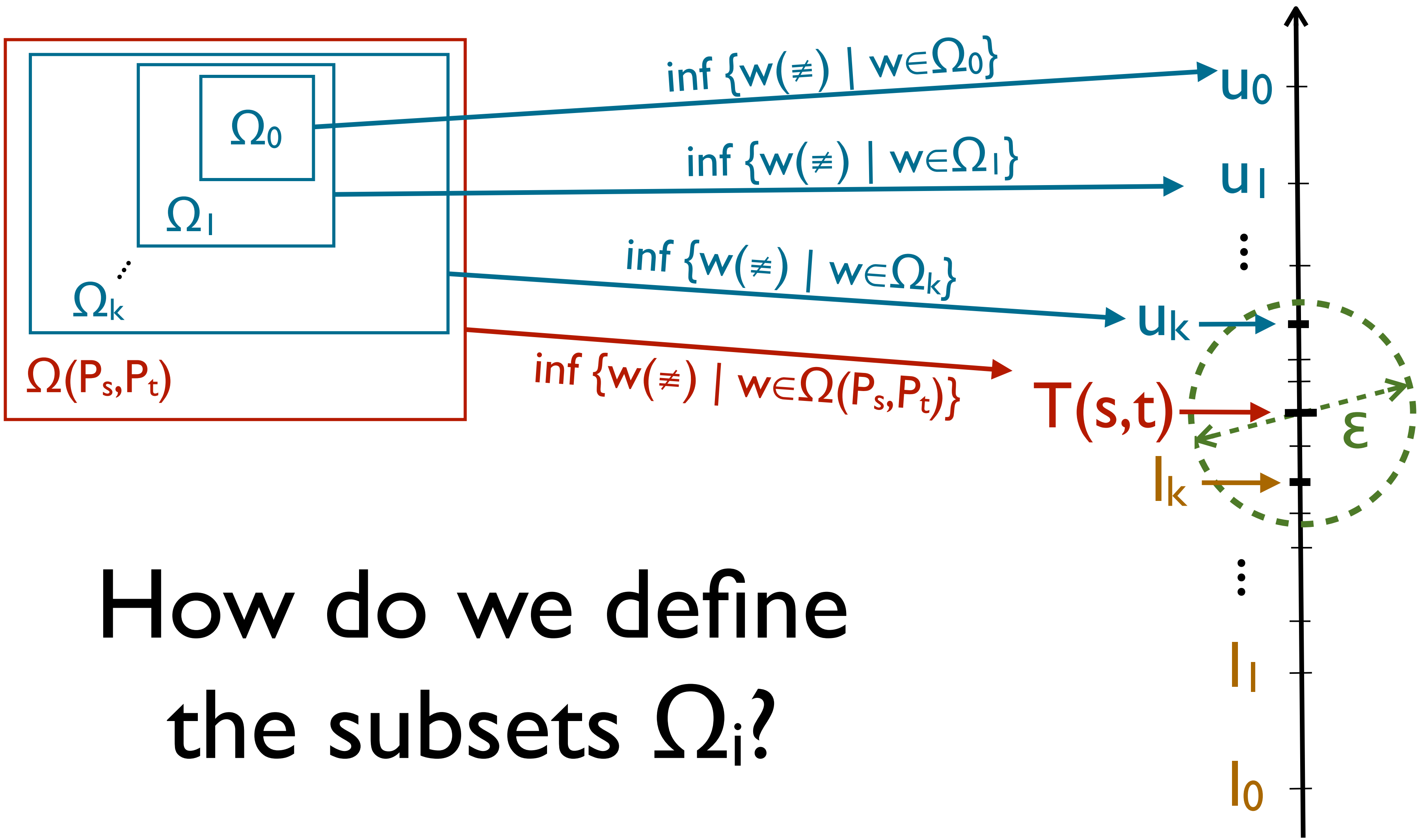
(general idea)

Approximation Schema



Approximation Schema

(general idea)



How do we define the subsets Ω_i?

Coupling Structure

Coupling Structure of rank k

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

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

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

Coupling Structure

Coupling Structure of rank k

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

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

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

Probability induced by $\mathcal{C}$ starting from (s,t)

$P_{\mathcal{C}}(s,t)$

Coupling Structure

Coupling Structure of rank k

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

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

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

Probability induced by $\mathcal{C}$ starting from (s,t)

$$\Omega_k = \{P_{\mathcal{C}}(s,t) \mid \mathcal{C} \text{ of rank } k\}$$

Coupling Structure

Coupling Structure of rank k

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

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

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

Probability induced by $\mathcal{C}$ starting from (s,t)

$$\Omega_k = \{ P_{\mathcal{C}}(s,t) \mid \mathcal{C} \text{ of rank } k \}$$

Lemma

(i) $\Omega_k \subseteq \Omega(P(s),P(t))$, (ii) $\Omega_k \subseteq \Omega_{hk}$ (for all $k,h>0$)

(iii) $\bigcup_k \Omega_k$ is dense in $\Omega(P(s),P(t))$

Computing the Approximants

(*) MC with rational transition probabilities

Computing the Approximants

- Both lower & upper approx. are computable

(*) MC with rational transition probabilities

Computing the Approximants

- Both lower & upper approx. are computable
- For each k>0, l_k and u_k can be computed in **polynomial time** in the size of the MC

(*) MC with rational transition probabilities

Computing the Approximants

- Both lower & upper approx. are computable
- For each k>0, l_k and u_k can be computed in **polynomial time** in the size of the MC

proved via alternative characterizations

(*) MC with rational transition probabilities

Computing the Approximants

- Both lower & upper approx. are computable
- For each k>0, l_k and u_k can be computed in **polynomial time** in the size of the MC

proved via alternative characterizations

the threshold problem for T(s,t) is still NP-hard!

(*) MC with rational transition probabilities

Upper approx. are Branching Metrics!

$$\Theta(d)(s,t) = \begin{cases} 1 & \text{if } s \neq t \\ K(d)(\tau(s),\tau(t)) & \text{otherwise} \end{cases}$$

Upper approx. are Branching Metrics!

$$\Theta(d)(s,t) = \begin{cases} 1 & \text{if } s \neq t \\ K(d)(\tau(s),\tau(t)) & \text{otherwise} \end{cases}$$

the 1st upper-approx is the least fixed point of the operator Θ

Upper approx. are Branching Metrics!

$$\Theta(d)(s,t) = \begin{cases} 1 & \text{if } s \neq t \\ K(d)(\tau(s),\tau(t)) & \text{otherwise} \end{cases}$$

Kantorovich lifting

the 1st upper-approx is the least fixed point of the operator Θ

Upper approx. are Branching Metrics!

$$\Theta(d)(s,t) = \begin{cases} 1 & \text{if } s \neq t \\ K(d)(\tau(s),\tau(t)) & \text{otherwise} \end{cases}$$

it is the Kantorovich distance of Desharnais at al.!

Kantorovich lifting

the 1st upper-approx is the least fixed point of the operator Θ

Upper approx. are Branching Metrics!

$$\Theta(d)(s,t) = \begin{cases} 1 & \text{if } s \neq t \\ K(d)(\tau(s),\tau(t)) & \text{otherwise} \end{cases}$$

Kantorovich lifting

it is the Kantorovich distance of Desharnais at al.!

the 1st upper-approx is the least fixed point of the operator Θ

its kernel is Larsen-Skou probabilistic bisimilarity!

Upper approx. are Branching Metrics!

$$\Theta^k(d)(s,t) = \begin{cases} 1 & \text{if } s \neq t \\ K(\Lambda^k(d))(\tau^k(s),\tau^k(t)) & \text{otherwise} \end{cases}$$

the k-th upper-approx is
the least fixed point
of the operator Θ^k

Upper approx. are Branching Metrics!

$$\Theta^k(d)(s,t) = \begin{cases} 1 & \text{if } s \neq t \\ K(\Lambda^k(d))(\tau^k(s), \tau^k(t)) & \text{otherwise} \end{cases}$$

the k-th upper-approx is the least fixed point of the operator Θ^k

its kernel is k-step generalization of probabilistic bisimilarity...

Upper approx. are Branching Metrics!

$$\Theta^k(d)(s,t) = \begin{cases} 1 & \text{if } s \neq t \\ K(\Lambda^k(d))(\tau^k(s),\tau^k(t)) & \text{otherwise} \end{cases}$$

k-steps transition

the k-th upper-approx is the least fixed point of the operator Θ^k

its kernel is k-step generalization of probabilistic bisimilarity...

Upper approx. are Branching Metrics!

$$\Theta^k(d)(s,t) = \begin{cases} 1 & \text{if } s \neq t \\ K(\Lambda^k(d))(\tau^k(s),\tau^k(t)) & \text{otherwise} \end{cases}$$

k-steps transition

Kantorovich lifting + ≢-selector

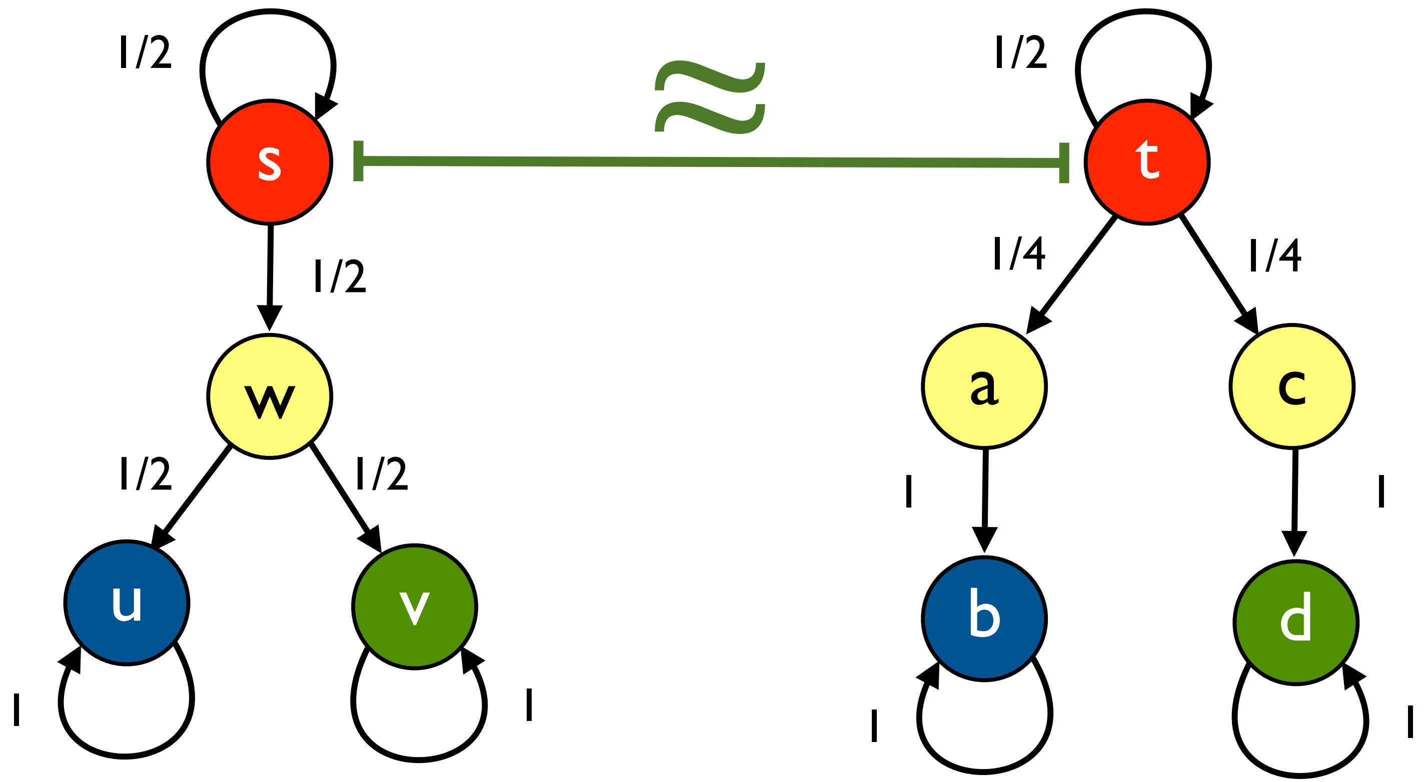
the k-th upper-approx is the least fixed point of the operator Θ^k

its kernel is k-step generalization of probabilistic bisimilarity...

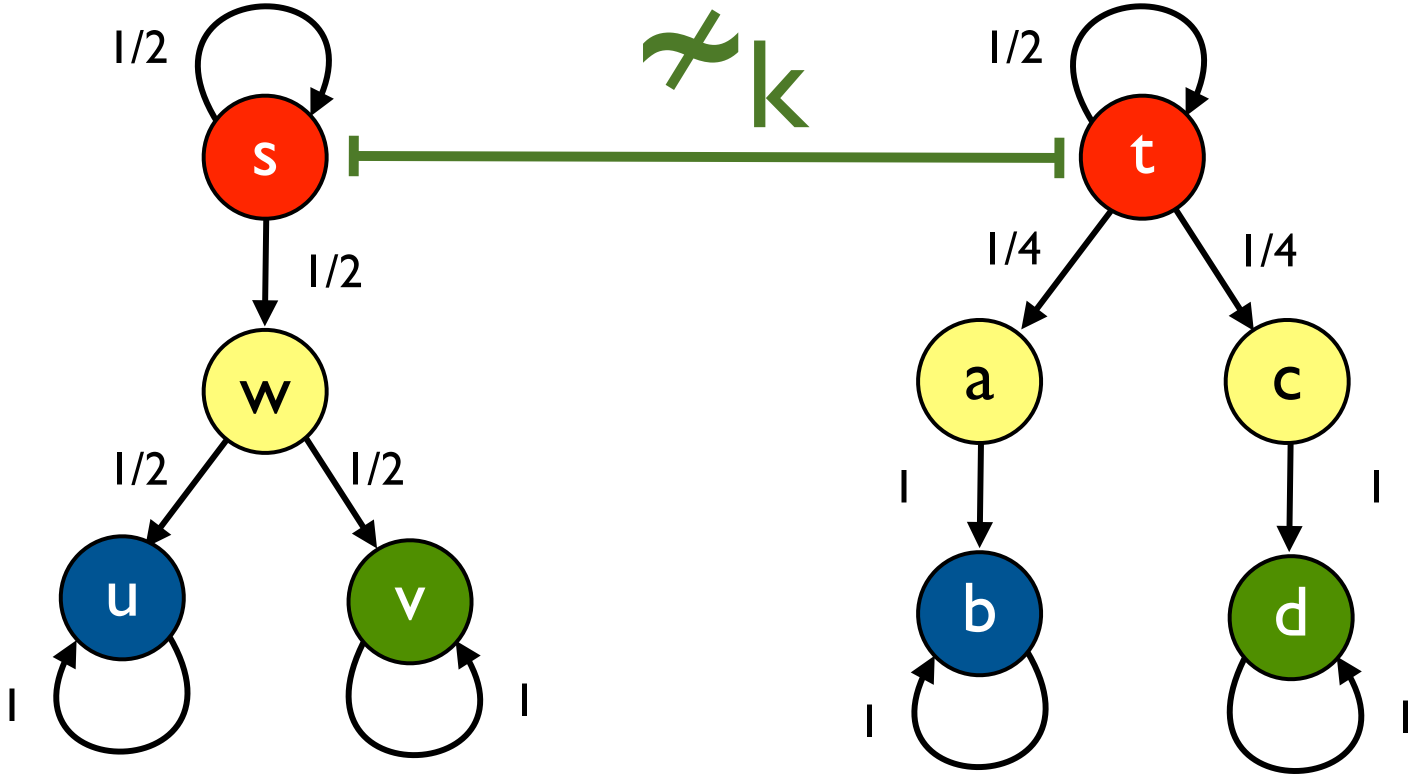
Exact semantics do NOT converge

	metric-based	equiv.-based
(monotone)	$u_k \geq u_{hk}$	$\sim_k \subseteq \sim_{hk}$
(bound)	$u_k \geq T$	$\sim_k \subseteq \approx$
(convergence)	$\inf_k u_k = T$	$\bigcup_k \sim_k \neq \approx$

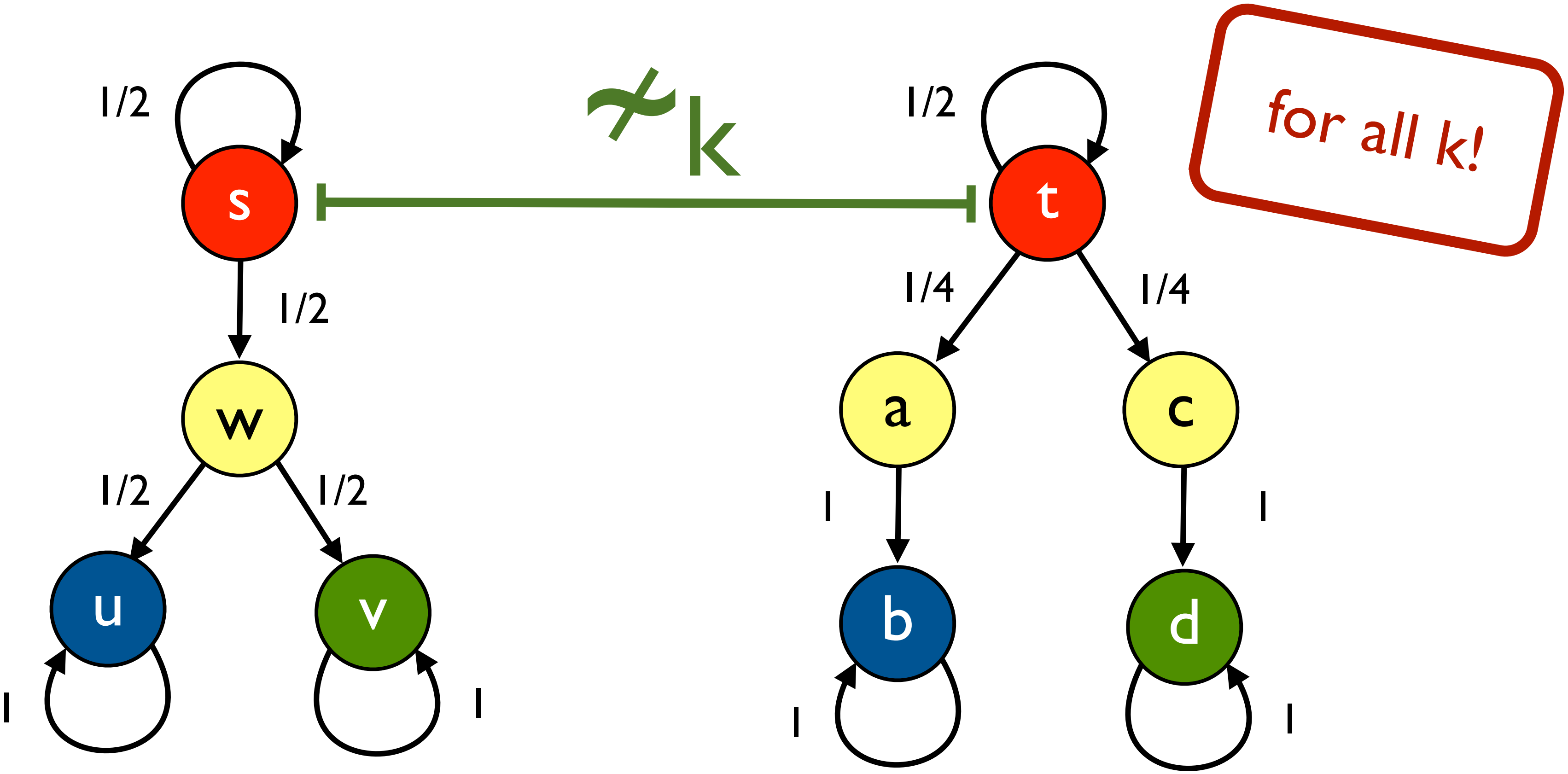
The Counterexample



The Counterexample



The Counterexample



Concluding Remarks

Concluding Remarks

- Metrics for Model Checking

Concluding Remarks

- Metrics for Model Checking
- Approximation algorithms (via duality)

Concluding Remarks

- Metrics for Model Checking
- Approximation algorithms (via duality)
- Branching converge to linear

Concluding Remarks

- Metrics for Model Checking
- Approximation algorithms (via duality)
- Branching converge to linear

Future Work

Concluding Remarks

- Metrics for Model Checking
- Approximation algorithms (via duality)
- Branching converge to linear

Future Work

- Better algorithms? (on-the-fly techniques)

Concluding Remarks

- Metrics for Model Checking
- Approximation algorithms (via duality)
- Branching converge to linear

Future Work

- Better algorithms? (on-the-fly techniques)
- different kind of models (non-determinism?)

Concluding Remarks

- Metrics for Model Checking
- Approximation algorithms (via duality)
- Branching converge to linear

Future Work

- Better algorithms? (on-the-fly techniques)
- different kind of models (non-determinism?)
- explore topological properties

Thank you
for the attention

Appendix

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 the supremum!