Converging from Branching to Linear Metrics for Weighted Transition Systems

Giorgio Bacci, Giovanni Bacci,
Kim G. Larsen, and Radu Mardare

Dept. Computer Science, Aalborg University

WATA 2016

Motivations

- Growing interest in ***quantitative aspects*** (probabilities, weights, time, etc.)
- ***Behaviors***: from equivalences to **distances**

value of the property φ at a given state

$$[\varphi] : S \longrightarrow \mathbb{R}$$

Equivalence

$$\forall \varphi.\ [\varphi](s) = [\varphi](t)$$

Distance

$$\sup_{\varphi} \left| [\varphi](s) - [\varphi](t) \right|$$

Linear vs Branching

Probabilistic case: (labelled Markov chains)

Linear-time	**Branching-time**
Trace Distance (a.k.a. total variation)	Bisimilarity Distance (a.k.a. Kantorovich)
Probabilistic LTL *(Bacci[2], Larsen & Mardare. 2015)*	**Probabilistic HML** *(Desharnais et al. 2004)*
NP-hard (undecidable?) *(Lyngsø & Pedersen et al. 2002)*	**Polynomial-time** *(Chen, van Breugel & Worrell. 2012)*

A converging sequence

nice

(Bacci², Larsen, Mardare. FoSSaCS15 & ICTAC15)

Trace distance — Bisimilarity distance

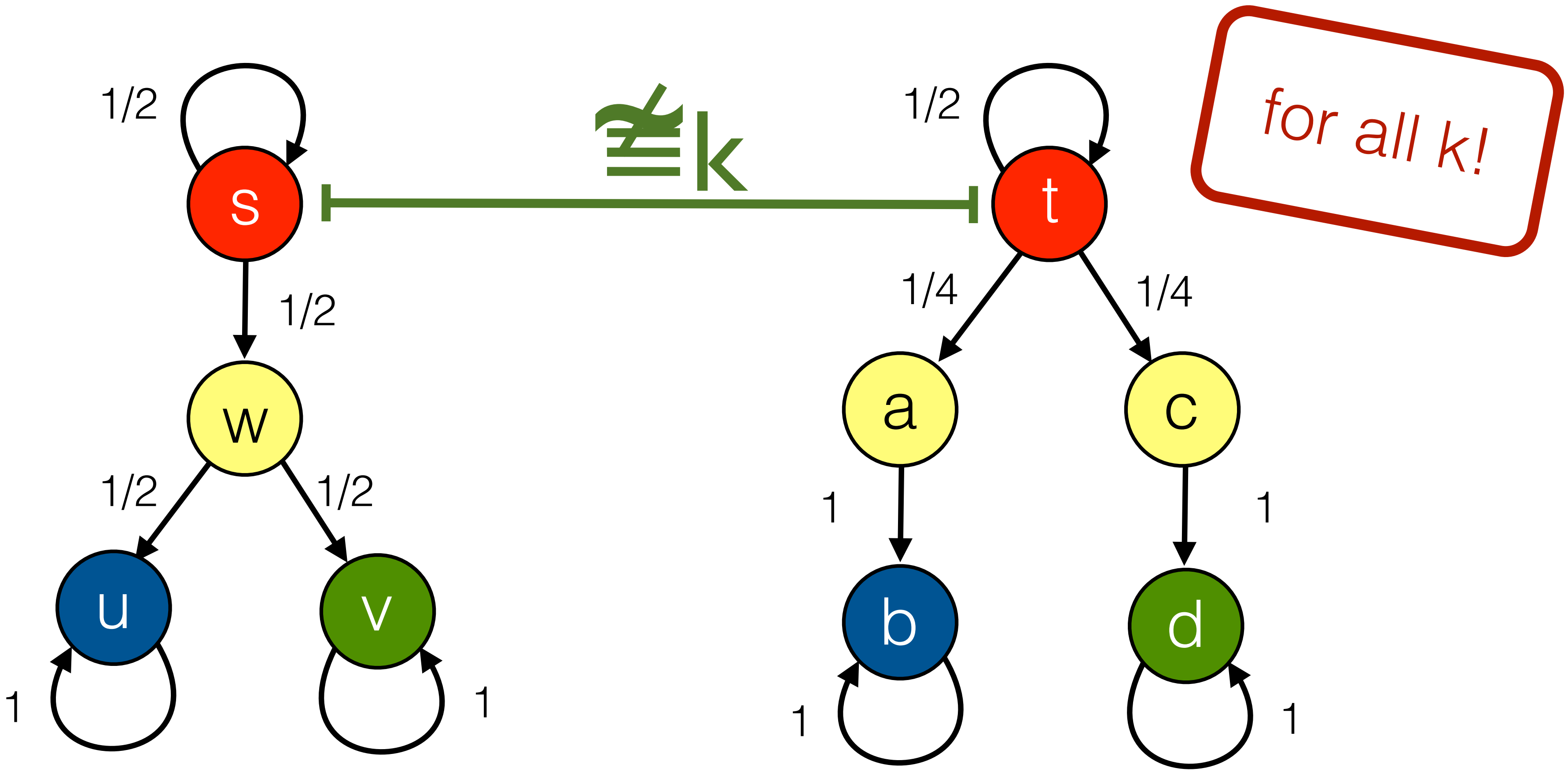
$$\mathbf{T} = \lim_k B_k \leftarrow \cdots \leq B_k \leq \cdots \leq B_3 \leq B_2 \leq B_1 = \mathbf{B}$$

$$\equiv \;\; \neq \;\; \bigcup_k \sim_k \leftarrow \cdots \supseteq \sim_k \supseteq \cdots \supseteq \sim_3 \supseteq \sim_2 \supseteq \sim_1 = \sim$$

Prob. trace equivalence — Probabilistic bisimilarity

Rough idea: "extending observations with a lookahead of k-steps"

Equivalences don't converge



$s \sim_k t$ iff for all $L_i \in S/{=_{\mathscr{L}}}$ and $C \in S/{\sim_k}$

$$P(s)(\mathfrak{C}(L_0..L_{k-1}C)) = P(t)(\mathfrak{C}(L_0..L_{k-1}C))$$

A converging sequence

nice

(Bacci et al. FoSSaCS15 & ICTAC15)

NP-hard

Polynomial-time

$$\mathbf{T} = \lim_k B_k \leftarrow \cdots \leq B_k \leq \cdots \leq B_3 \leq B_2 \leq B_1 = \mathbf{B}$$

The ingredients for the convergence

1. **Kantorovich Duality** (*1 table spoon*)
2. Coupling Structures (∞*-many for each k*)
3. Density / Saturation argument (*mix until smooth*)

The recipe comes in different flavors

Examples:

Labelled Markov Chains

Weighted Transition Systems

in parallel

Labelled Markov Chains

$$\mathcal{M} = (S,\ \tau\colon S \to \Delta(S),\ l\colon S \to \mathscr{L})$$

set of states

transition probability

labelling function

Trace distance

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

prob. on ω-traces (up to trace equivalence)

Bisimilarity distance

$$B(s,t) \overset{\text{lfp}}{=} \max\{\mathbb{1}_{\mathscr{L}}(l(s), l(t)),\ K(B)(\tau(s), \tau(t))\}$$

Kantorovich distance

Weighted Transition Systems

$\mathscr{W} = (S,\ \theta: S \to \mathscr{P}(S),\ w: S \to \mathbb{K})$ — ($\mathbb{K}$,d) metric space

- set of states
- transition function
- weight function

PSPACE-complete

point wise Trace distance

$$T(s,t) = H(d^{\omega})(Tr(s),Tr(t))$$

Hausdorff distance

set of ω-traces

P-time

point wise Branching distance

$$B(s,t) \overset{lfp}{=} \max\{d(w(s),w(t)),\ H(B)(\theta(s),\theta(t))\}$$

Hausdorff distance

(de Alfaro et al. 2009)

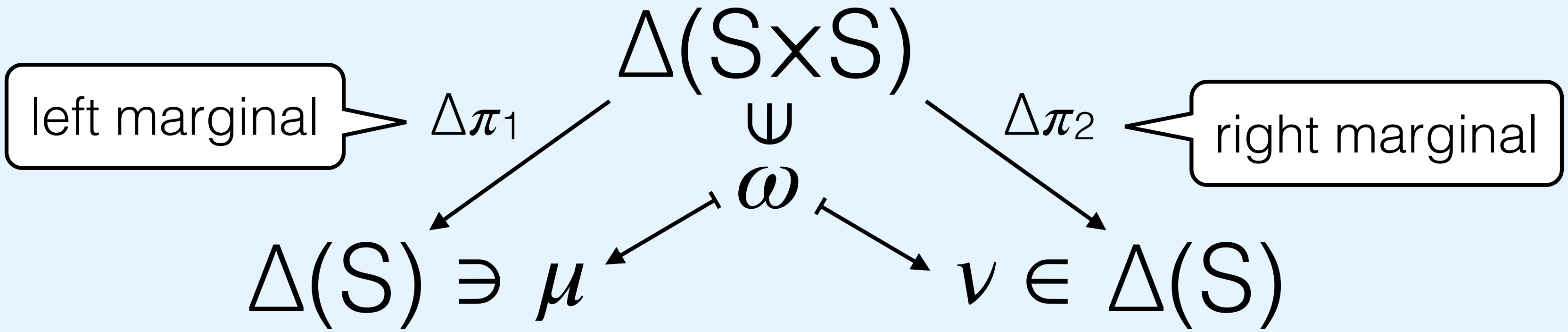
1st ingredient

DUALITY

Kantorovich duality

$$K(d)(\mu,\nu) = \sup \left\{ \left| \int f \, \mathbf{d}\mu - \int f \, \mathbf{d}\nu \right| : f \in \mathrm{Nexp} \right\}$$

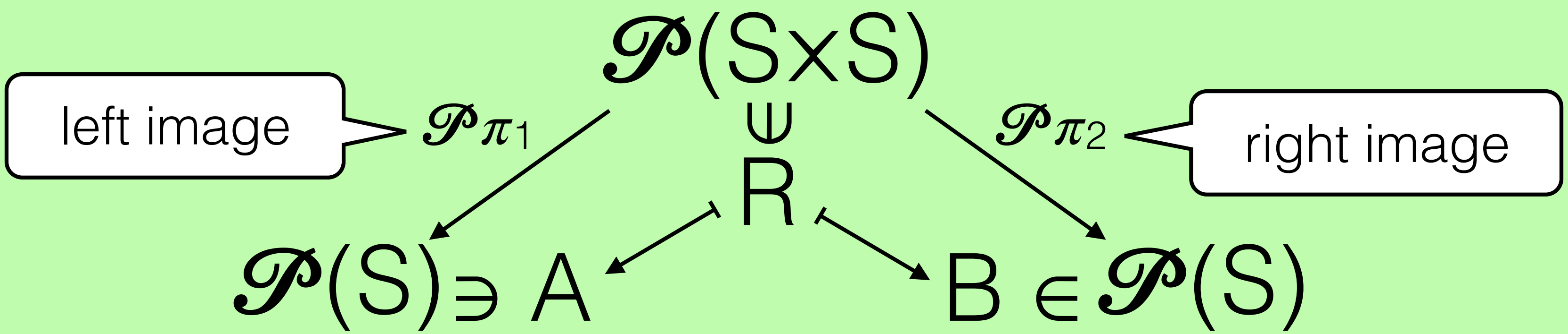
$$= \min \left\{ \int d \, \mathbf{d}\omega : \omega \in \Omega(\mu,\nu) \right\}$$



Hausdorff duality

$$H(d)(A,B) = \max\left\{\sup_{a\in A} d(a,B),\ \sup_{b\in B} d(b,A)\right\}$$

$$= \inf\left\{\sup_{(x,y)\in R} d(x,y) : R \in \mathbf{\Gamma}(A,B)\right\}$$



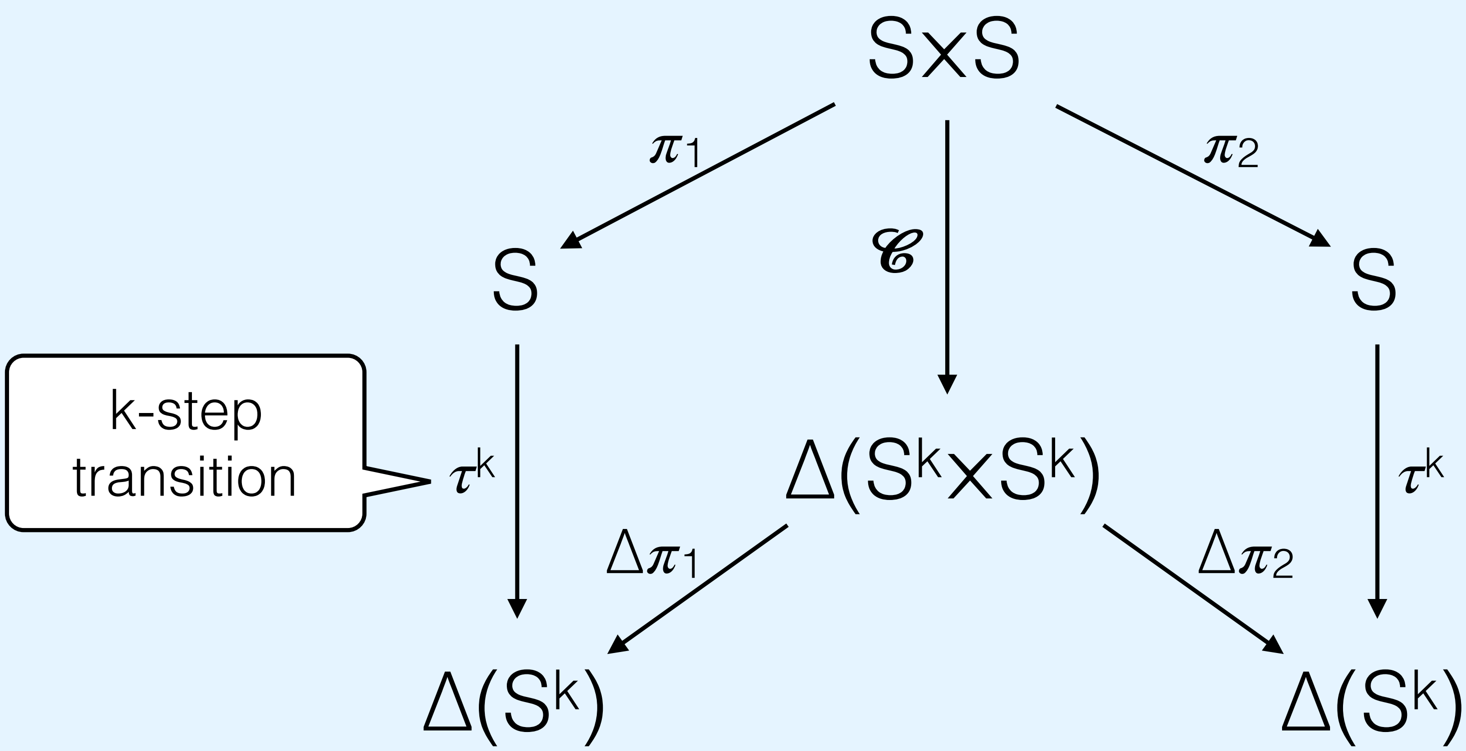
2nd ingredient

Coupling Structures

Coupling Structure of rank k

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

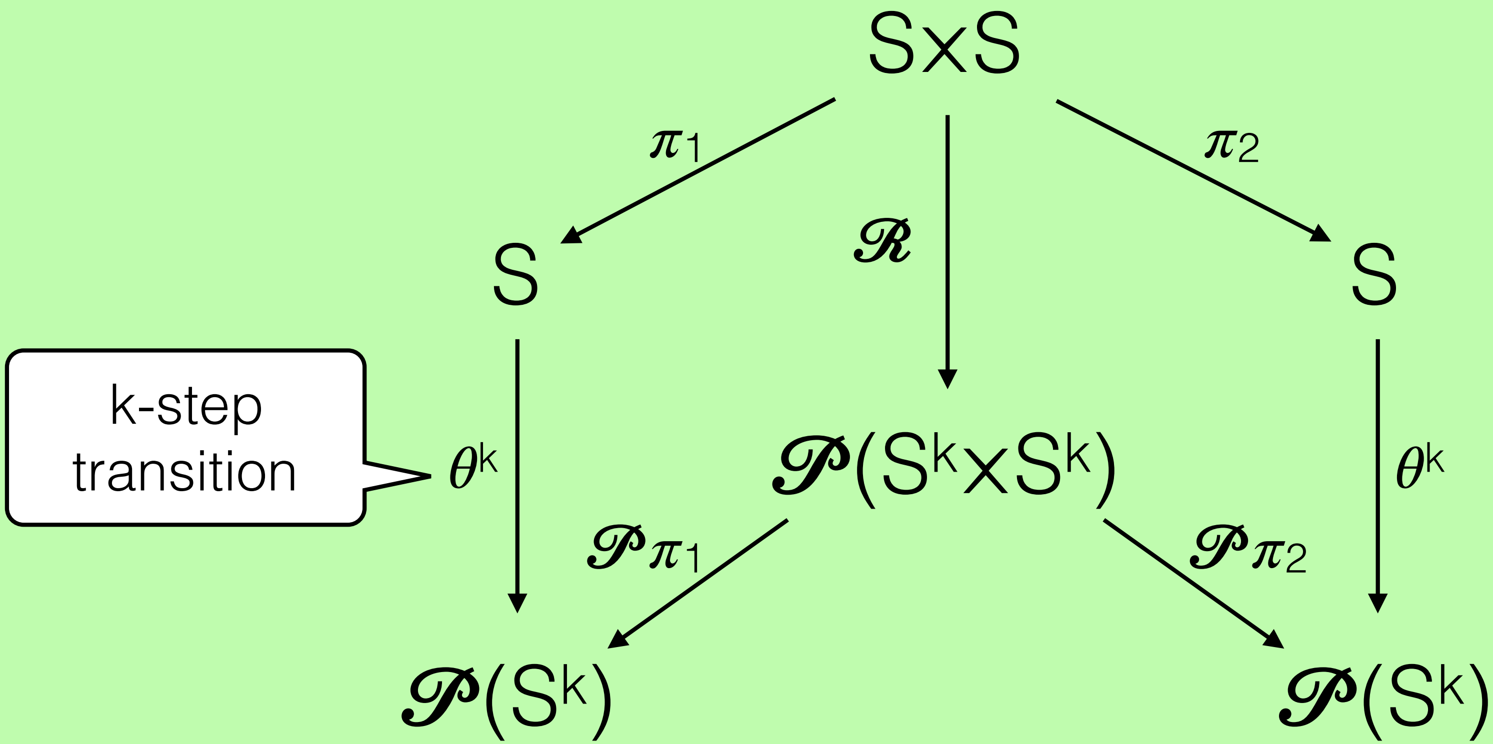
such that $\mathscr{C}(s,t) \in \Omega(\tau^k(s), \tau^k(t))$



Rel. Coupling Struct. of rank k

$$\mathscr{R}: S\times S \to \mathscr{P}(S^k \times S^k)$$

such that $\mathscr{R}(s,t)\in\boldsymbol{\Gamma}(\theta^k(s),\theta^k(t))$



From coupling structures to couplings on ω-traces

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

$$\Downarrow$$

$$P^{\mathscr{C}} : S \times S \to \Delta(S^\omega \times S^\omega)$$

$P^{\mathscr{C}}(s,t)$ is the probability induced by $\mathscr{C}$ over pairs of ω-traces

Lemma

$P^{\mathscr{C}}(s,t) \in \Omega(P(s),P(t))$

$$\Omega_k(s,t) = \{P^{\mathscr{C}}(s,t) \mid \mathscr{C} \text{ coupl. struct. of rank } k\}$$

From rel. coupling structures to rel. couplings on ω-traces

$$\mathscr{R}: S\times S \to \mathscr{P}(S^k \times S^k)$$

$$\downarrow$$

$$Q^{\mathscr{R}}: S\times S \to \mathscr{P}(S^\omega \times S^\omega)$$

$Q^{\mathscr{R}}(s,t)$ is the relation over pairs of ω-traces induced by $\mathscr{R}$ via composition

Lemma

$$Q^{\mathscr{R}}(s,t) \in \mathbf{\Gamma}(Tr(s), Tr(t))$$

$$\mathbf{\Gamma}_k(s,t) = \{Q^{\mathscr{R}}(s,t) \mid \mathscr{R} \text{ coupl. struct. of rank } k\}$$

Coupling Characteriza

duality

Trace distance

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

k-Bisimilarity distance

$$B_k(s,t) = \min\{\omega(\neq) \mid \omega \in \Omega_k(s,t)\}$$

Lemma

(i) $\Omega_k \subseteq \Omega(P(s),P(t))$ and (ii) $\Omega_k \subseteq \Omega_{k+1}$

hence:

$$\mathbf{T} \leq \lim_k B_k \leftarrow \cdots \leq B_k \leq \cdots \leq B_2 \leq B_1 = \mathbf{B}$$

Coupling Characteriza[tion]

duality

point wise Trace distance

$$T(s,t) = \inf \{ \sup_{(\pi,\rho)\in R} d^{\omega}(\pi,\rho) \mid R \in \mathbf{\Gamma}(Tr(s),Tr(t)) \}$$

k-point wise Branching distance

$$B_k(s,t) = \inf \{ \sup_{(\pi,\rho)\in R} d^{\omega}(\pi,\rho) \mid R \in \mathbf{\Gamma}_k(s,t) \}$$

Lemma

(i) $\mathbf{\Gamma}_k \subseteq \mathbf{\Gamma}(Tr(s),Tr(t))$ and (ii) $\mathbf{\Gamma}_k \subseteq \mathbf{\Gamma}_{k+1}$

hence:

$$\mathbf{T} \leq \lim_k B_k \leftarrow \cdots \leq B_k \leq \cdots \leq B_2 \leq B_1 = \mathbf{B}$$

3rd ingredient

Density/Saturation

Density/Saturation

Trace distance

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

k-Bisimilarity distance

$$B_k(s,t) = \min \{\omega(\neq) \mid \omega \in \Omega_k(s,t) \}$$

Lemma

(i) $U_k\Omega_k \neq \Omega(P(s),P(t))$

(ii) $U_k\Omega_k \overset{\textbf{dense}}{\subseteq} \Omega(P(s),P(t))$

hence:

$$T = \lim_k B_k$$

Density/Saturation

point wise Trace distance

$$T(s,t) = \min \{ \sup_{(\pi,\rho)\in R} d^{\omega}(\pi,\rho) \mid R \in \boldsymbol{\Gamma}(Tr(s),Tr(t)) \}$$

k-point wise Branching distance

$$B_k(s,t) = \min \{ \sup_{(\pi,\rho)\in R} d^{\omega}(\pi,\rho) \mid R \in \boldsymbol{\Gamma}_k(s,t) \}$$

Lemma

(i) $\bigcup_k \boldsymbol{\Gamma}_k \neq \boldsymbol{\Gamma}(Tr(s),Tr(t))$

(ii) $\bigcup_k \boldsymbol{\Gamma}_k \overset{\textbf{dense}}{\subseteq} \boldsymbol{\Gamma}(Tr(s),Tr(t))$

hence:

$$T = \lim_k B_k$$

Open Questions

- To what extent can this recipe be **generalized**?
  (via functor Lifting… as in Baldan et al. 2014)

- Is this construction **compositional**?
  (composition of behavior functors)

- Can this convergence be exploited for
  faster approximations of the linear distances?

Thank you
for your attention