Approximate Simulations for Probabilistic I/O Automata

Sayan Mitra and Nancy Lynch

TDS Seminar, CSAIL, MIT

October 27th 2006

Implementation Relations

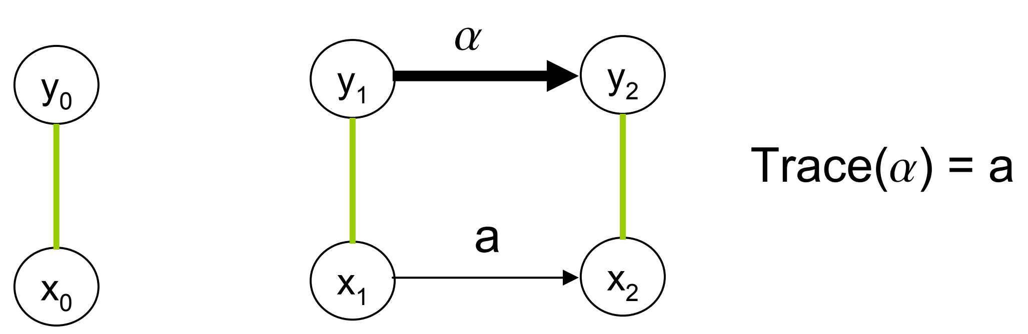
- Timed automata, probabilistic automata
- Executions record evolution of the system
- Traces or visible behaviors
  - sequences of external actions
  - sequences of external actions and intervals
  - probability distribution over visible sequences
- Implementation as trace inclusion (e.g., Lynch, Vaandrager 1995)
- A implements B if Traces(A) $\subseteq$ Traces(B), written as A $\leq$ B
- A and B are equivalent if Traces(A) = Traces(B), written as A = B

- Other notions of implementation : strong and weak bisimulation,

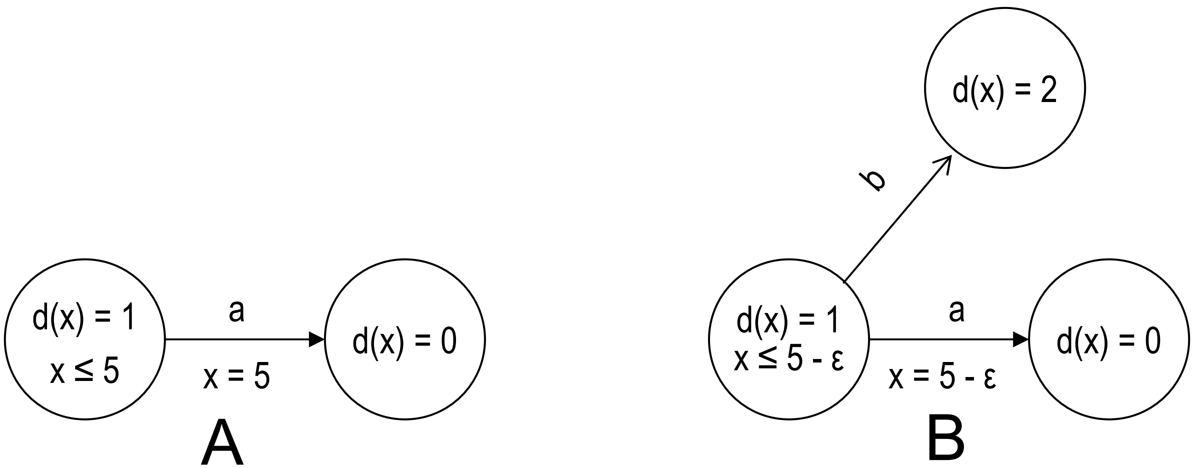
  reachable set inclusion

Simulation Relations for Proving Implementation

- A, B automata
- R is a relation on X×Y
- Forward simulation

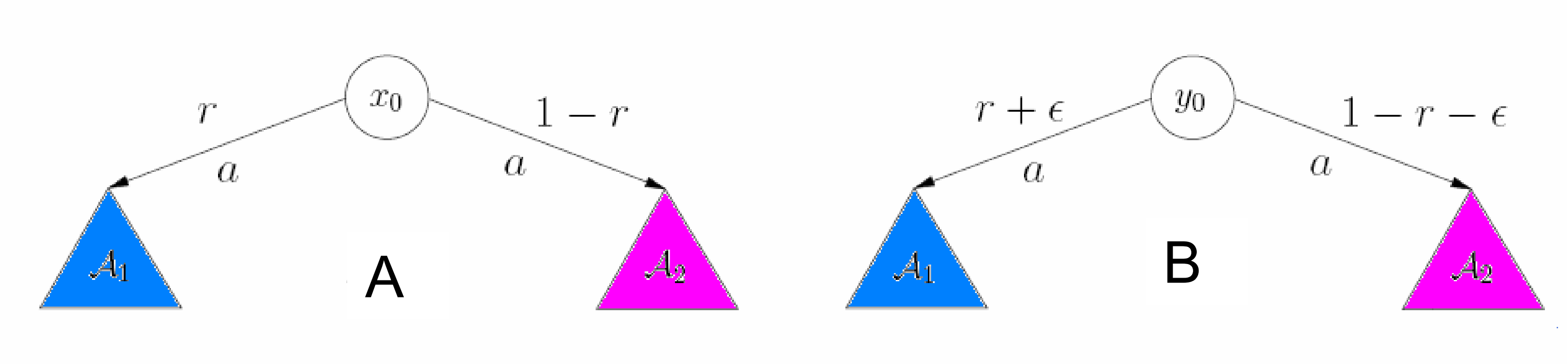


Fragility of Classical Implementation in Timed Systems



- Traces(A) = { [0,t],[0,5]a,…}
- Traces(B) = { [0,t],[0,5+ε]a,…,[0,t]b,…}
- A and B cannot be compared

and in Probabilistic Systems



- Suppose (α_1,p_1) and (α_2,p_2) are the only traces of A_1 and A_2
- Traces(A) contains $(a\alpha_1, rp_1)$ $(a\alpha_2,(1-r)p_2)$
- Traces(B) contains $(a\alpha_1,(r+\varepsilon)p_1)$ $(a\alpha_2,(1-r-\varepsilon)p_2)$
- Again, A B cannot be compared

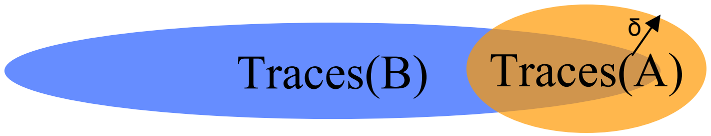
Metrics on Traces

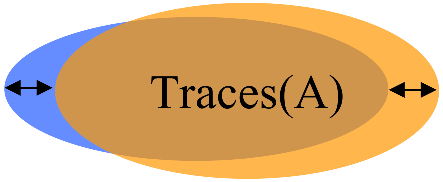
Define a metric **d** on the space **T** containing traces of A and traces of B

(**T**,**d**) is a metric space

A *δ-implements* ***B*** if for every μ_A in Trace(A) there is μ_B in Trace(B) s.t. **d** $(\mu_A,\mu_B) \leq \delta$

A and **B** are *δ-equivalent* if they δ-implement each other

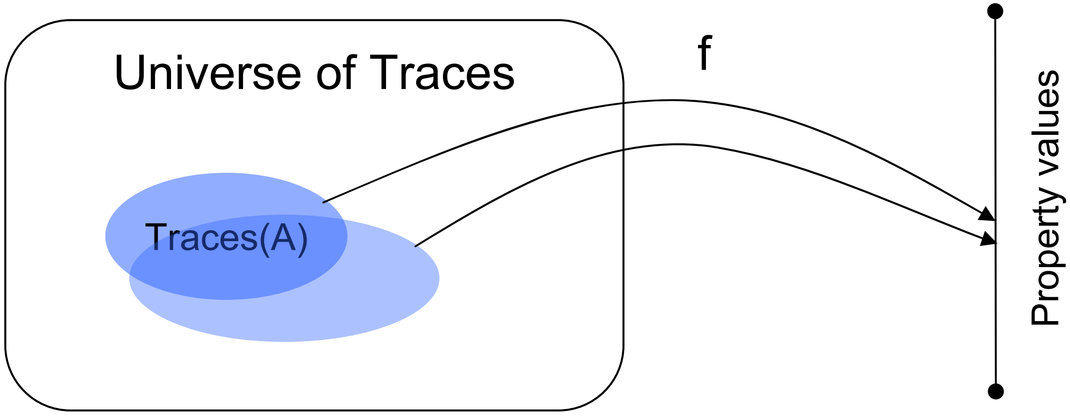




Metrics on Traces

- Robust implementation relations
- Actual value of δ is unimportant
- Abstraction and model reduction through approximation

- Continuity of quantitative properties



- Define metrics for trace distributions of PIOA
- Simulation based techniques for proving **A** *δ-implements* ***B***

Outline

- Task-structured PIOA
- A simple approximate simulation
- Expanded approx. sim.
- Discounted approx. sim.
- Other metrics and new directions

Related Work I

1. Probabilistic CCS
   Giacalone, Jou, Smolka (1994)
   Introduced metric in the study of bisimilarity
   Metric: *ε-bisimilar*

2. Probabilistic Concurrent 2-player Games
   Alfaro, Henzinger, Mazumdar (2003)
   Discounted mu-calculus
   Fixpoint-based algorithms for checking discounted properties

Related Work II

3. Labelled Markov Processes (LMPs)
   Desharnais, Gupta, Jagadeesan, Panagaden (2002-04)
   - Metric on states defined based on a class
   - d(A,B) = 0 implies A and B are bisimilar

   Van Breugel, Mislove, Ouaknine, Worell (2003)
   - Intrinsic characterization of the above metric
   - Topology induced by the above metric on the space of LMPs
   - Polynomial time algorithm for the metric for finite LMPs

4. Generalized Semi Markov Processes
   Gupta, Jagadeesan, Panagaden (2004)
   - Pseudometric analogue of bisimulation
   - Continuity of properties

No nondeterminism, based on bisimulations

Related Work III

5. Discrete/Continuous Metric Transition Systems

   Girard, Pappas (2005)

   - Pseudometrics on: trace inclusion, Reachable set inclusion, bisimulation
   - Algorithms for computing metrics

6. Probabilistic I/O Automata

   Cheung (2006)

   - Trace distributions are closed sets in $([0,1]^{Traces^*},d)$
   - Finite tests are sufficient to distinguish infinite processes

Task-Structured PTIOA

Canetti, Cheung, Kaynar, Liskov, Lynch, Pereira, Segala(2005-06)

Discrete probability measures on X, $\mu(E) = \Sigma_{e \in E} \mu(\{e\})$ and $\mu(X) = 1$

Disc(X), Supp(μ)

A = *(Q,v,I,O,H,D,R)*

Q: countable set of states

v: initial distribution

I,O, H: countable, pairwise disjoint sets of actions

D $\subseteq$ Q × *A* × *Disc(Q),* (q,a,μ) ϵ D is written as $q \rightarrow_a \mu$

R: Equivalence relation on L; each equivalence class of *R* is a *task (T)*

Axioms:

- Input enabled
- For any q & a, there is at most one $q \rightarrow_a \mu$
- For any q & T, there is at most one a ϵ T enabled at q.

For this talk assume A is closed

Executions and Traces

As usual

- Execution fragment $\alpha = q_0 a_1 q_1 a_2 \ldots$
- α is an execution if q_0 in Supp(v)
- Execs, Execs*
- trace(α) captures the visible part of α
  - delete all q's and the a's in H
- Traces, Traces*

But PIOA is probabilistic

"visible behavior" = distribution over Traces, *a trace distribution*

Nondeterministic, therefore set of trace distributions

Canetti, Cheung, Kaynar, Liskov, Lynch, Pereira, Segala(05, 06)

Task Scheduler

- *Task scheduler* for A is a (finite or infinite) sequence of tasks T_1 T_2…
  - It interacts with A to give discrete distributions over execution fragments
- For this talk assume task scheduler is finite
  - All distributions are *finite* (we avoid limit arguments)
  - A finite measure can be viewed as a discrete measure on finite fragments

In general

- σ-field on Execs generated by cones
- discrete σ-field of Execs* is contained in the above
- Likewise σ-fields for Traces
- Construct chains of measures and then take limits

Applying a Schedule

Given a distribution μ over Execs* a task schedule $\sigma = T_1 T_2 \ldots T_n$
apply(μ,σ) gives a probability distribution Execs by applying σ

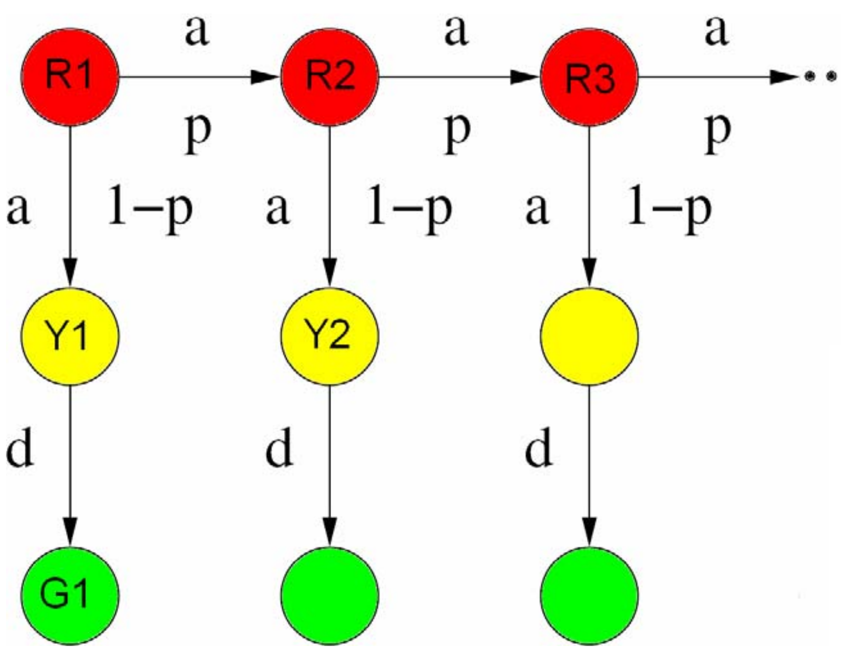
- apply(μ,⊥) = μ
- apply(μ,T) = μ'
- apply(μ, σT) = apply(apply(μ, σ), T)

- apply(μ, σ) = $\lim_{i \to \infty}$ apply(μ, σ_i), σ_i is the length i prefix of σ

Canetti, Cheung, Kaynar, Liskov, Lynch, Pereira, Segala(05, 06)

- apply(μ,σ) = probability distribution over fragments
  - apply(μ,$\perp$) = μ
  - apply(μ,T) = μ'
    - $\mu'(\alpha) = p_1(\alpha) + p_2(\alpha)$
    - $p_1(\alpha) = \mu(\alpha')\eta(q)$ if $\alpha = \alpha'$aq and a is in task T and lstate(α')$\rightarrow_a \eta$
    - $p_2(\alpha) = \mu(\alpha)$ if T is *not* enabled in lstate(α)
  - apply(μ, σT) = apply(apply(μ, σ), T)

Applying Schedules



α / σ										
$\perp$	1									
a		1-p	p							

$p_1(\alpha) = \mu(\alpha')p$ if $\alpha = \alpha' a\, R_{i+1}$ lstate$(\alpha') = R_i$

$= \mu(\alpha')(1\text{-}p)$ if $\alpha = \alpha' a\, Y_i$ lstate$(\alpha') = R_i$

$= \mu(\alpha')$ if $\alpha = \alpha' d\, G_i$ lstate$(\alpha') = Y_i$

$p_2(\alpha) = \mu(\alpha)$ if lstate$(\alpha) = R_i$

$= \mu(\alpha)$ if lstate$(\alpha) = Y_i$

- A task schedule σ defines μ=apply(v, σ) is a *probabilistic execution*

- Corresponding *trace distribution* tdist(μ)(β) = μ($trace^{-1}$(β))

- tdists(A) = {tdist(apply(v, σ)): σ is a task scheduler for A}
  set of all possible trace distributions

- A≤B if tdists(A) $\subseteq$ tdists(B)

Metrics on Trace Distributions

A metric $d : Disc(E^*) \times Disc(E^*) \rightarrow R$

(1) $d(\mu_1, \mu_2) = 0$ iff $\mu_1 = \mu_2$

(2) $d(\mu_1, \mu_2) = d(\mu_2, \mu_1)$

(3) $d(\mu_1, \mu_3) \leq d(\mu_1, \mu_2) + d(\mu_2, \mu_3)$

A δ-*implements* B (w.r.t metric d) if for every trace dist μ_1 of A there is a trace dist μ_2 of B such that $d(\mu_1,\mu_2) \leq \delta$.

We write this as $A \leq_\delta B$.

A and B are δ-equivalent if they δ-*implement* each other.

We write this as $A =_\delta B$.

Simple Approximate Simulation Relation (SA)

Given $\varepsilon, \delta > 0$, a function $\phi : Disc(Exec^*(A)) \times Disc(Exec^*(B)) \rightarrow \Re_{\geq 0} \cup \{\infty\}$

is a Simple Approximate (SA) Simulation function if

1. Start : $\phi(v_1, v_2) \leq \varepsilon$

2. Step : $\phi(\mu_1, \mu_2) \leq \varepsilon$ implies $\phi(\mu_1', \mu_2') \leq \varepsilon$

3. Trace : $\phi(\mu_1, \mu_2) \leq \varepsilon$ implies $\mathbf{d}(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta$

Segala (1995-96)

Simulation Relation:

$R \subseteq$ Disc(Execs*(A)) × Disc(Execs*(B))

1. $v_1 R v_2$
2. $\mu_1 R\ \mu_2$, implies $\mu_1'\ E(R)\ \mu_2'$
3. $\mu_1 R\ \mu_2$ implies $\text{tdist}(\mu_1) = \text{tdist}(\mu_2)$

Simple Approximate Simulation Relation (SA)

Given $\varepsilon, \delta > 0$, a function $\phi : Disc(Exec^*(A)) \times Disc(Exec^*(B)) \rightarrow \Re_{\geq 0} \cup \{\infty\}$

is a Simple Approximate (SA) Simulation function if

1. Start : $\phi(v_1, v_2) \leq \varepsilon$

2. Step : There exists a function $c : R_1^* \times R_1 \rightarrow R_2^*$ such that,

if $\phi(\mu_1, \mu_2) \leq \varepsilon$,

σ is a task schedule and T is a task for A,

and μ_1 is consistent with σ

and μ_2 is consistent with $full(c)(\sigma)$

then $\phi(apply(\mu_1, T), apply(\mu_2, c(\sigma, T))) \leq \varepsilon$

3. Trace : $\phi(\mu_1, \mu_2) \leq \varepsilon$ implies $\mathbf{d_u}(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta$

Soundness of SA

Theorem 1. If there exists an (ε,δ)-SA simulation function from A to B then $A \leq_\delta B$.

- Consider any $\mu_A = \text{apply}(v_1, T_1 T_2 \ldots T_n)$

  $\sigma_j = c(T_1 \ldots T_j)$

  $\mu_{A,j} = \text{apply}(v_1, T_1 T_2 \ldots T_j)$

  $\mu_{B,j} = \text{apply}(v_1, \sigma_1 \sigma_2 \ldots \sigma_j)$

  $\mu_B = \mu_{B,n}$
- For all j, $\phi(\mu_{A,j}, \mu_{B,j}) \leq \varepsilon$ (by induction using 1,2)
- For all j, $d(\text{tdist}(\mu_{A,j}), \text{tdist}(\mu_{B,j})) \leq \delta$ (by 3)

In particular, $d(\text{tdist}(\mu_A), \text{tdist}(\mu_B)) \leq \delta$

- For infinite task schedules take limit of a sequence of probability measures.

$$\lim_{j\to\infty}\eta_{Aj}=\eta_A \text{ and } \lim_{j\to\infty}\eta_{Bj}=\eta_B \text{ then } \lim_{j\to\infty} \mathrm{d}(\eta_{Aj},\eta_{Bj})=\mathrm{d}(\eta_A,\eta_B)$$

- **Step** and **Trace** conditions are critical for the choice of the metric and the simulation function

Probabilistic Safety

$$\mathbf{d}_u(\mu_1, \mu_2) = \sup_{C \subseteq 2^{Traces*}} | \mu_1(C) - \mu_2(C) |$$

Define a function (random variable) **X**:Traces →{0,1}

X(β) := 1 if some bad action occurs in β
0 otherwise

Suppose A is safe with probability at least p and B$\leq_\delta$ A

Claim: B is safe with probability at least δ + p

Let μ_B be any trace distribution of B.
There exits μ_A such that, for all C, | μ_B(C) - μ_A(C) | ≤ δ.
Then, μ_B([**X**=1]) ≤ δ + μ_A([**X**=1]) ≤ δ + p

SA Simulation

- A simulation relation R tells us if two distributions are related or not
- An approximate simulation function ϕ gives us a measure of how close two distributions are
- We want to capture internal branching

- So, decompose μ_A, μ_B into close components
  - $\mu_A(\alpha) = \Sigma_i \lambda_{Ai} \mu_{Ai}(\alpha)$, where $\Sigma_i \lambda_{Ai} = 1$
  - $\mu_B(\alpha) = \Sigma_i \lambda_{Bi} \mu_{Bi}(\alpha)$, where $\Sigma_i \lambda_{Bi} = 1$
  - $\phi(\mu_{Ai}, \mu_{Bj})$ is small

- More powerful simulation function

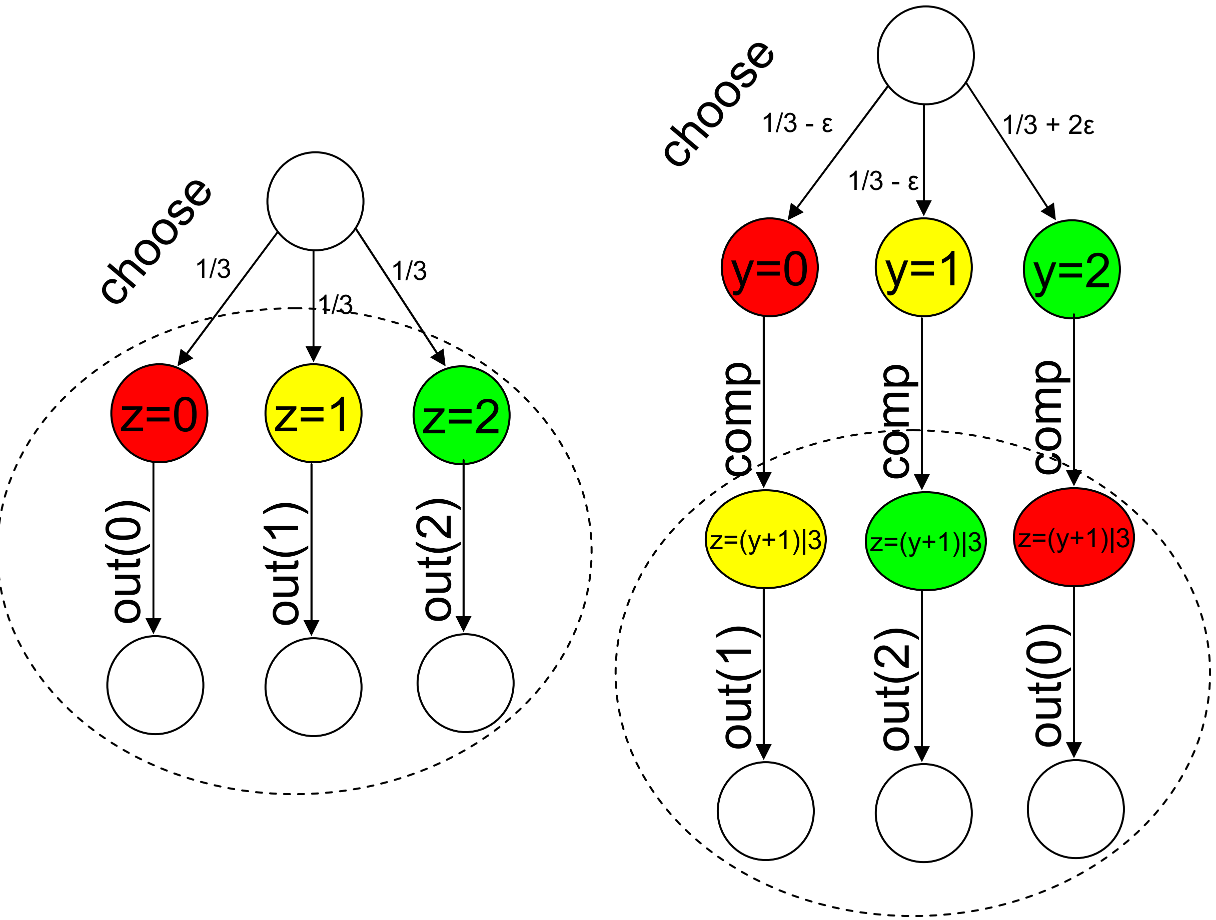
SA Simulation

- A simulation relation R tells us if two distributions are related or not
- An approximate simulation function ϕ gives us a measure of how close two distributions are
- We want to capture internal branching

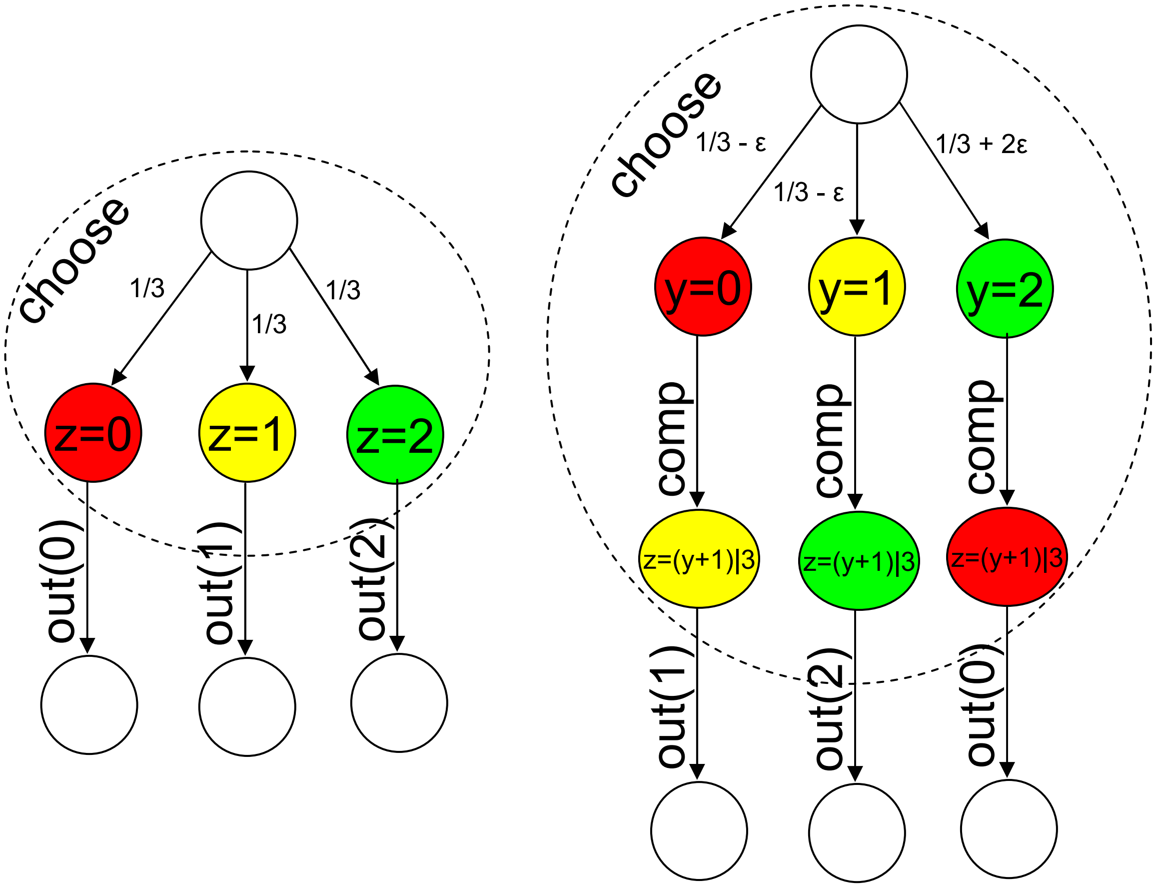
- So, decompose μ_A, μ_B into close components
  - $\mu_A(\alpha) = \Sigma_i \, \eta_A(\mu_{Ai}) \, \mu_{Ai}(\alpha)$, where η_A is Disc(Disc(Exec*$_A$))
  - $\mu_B(\alpha) = \Sigma_i \, \eta_B(\mu_{Bi}) \, \mu_{Bi}(\alpha)$, where η_B is Disc(Disc(Exec*$_B$))
  - $\phi(\mu_{Ai}, \mu_{Bj})$ is small

- More powerful simulation function

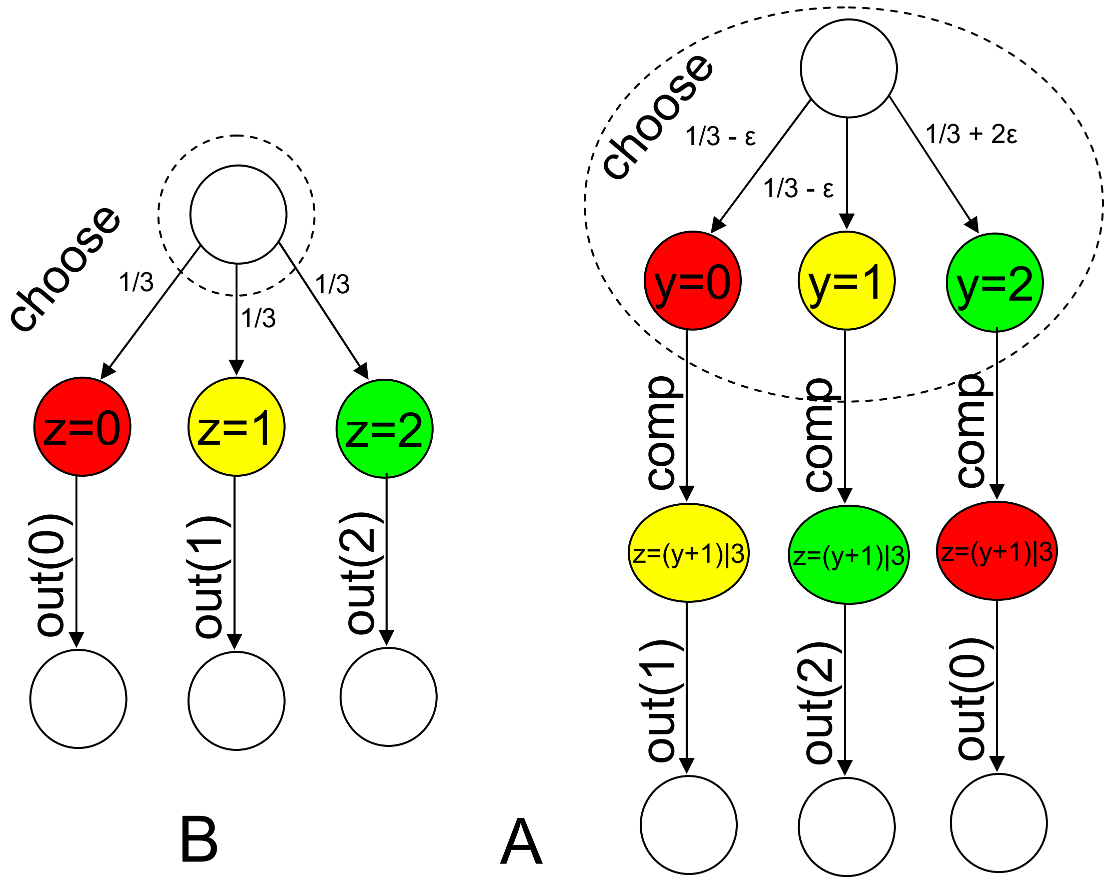
Need for Expansion



Need for Expansion



Need for Expansion



s, u : support of lsate(μ_A), lstate(μ_B)

$$\phi(\mu_A, \mu_B) = \begin{cases} \max_{\alpha,\beta} \mu_A(\alpha) + \mu_B(\beta) & \text{if } \exists \text{s, u, s.z} \neq \text{u.z} \\ 0 & \text{if } \forall \text{s, u, s.z} = \text{u.z} = \bot \text{ and s.y} = \bot \\ \max_{\alpha} \left| \mu_A(\alpha) - 1/3 \right| & \text{otherwise} \end{cases}$$

$$\phi(\mu_A,\mu_B)=\begin{cases}\max_{\alpha,\beta}\mu_A(\alpha)+\mu_B(\beta) & \text{if } \exists s,u,\ s.z\neq u.z\\ 0 & \text{if } \forall s,u,\ s.z=u.z=\perp \text{ and } s.y=\perp\\ \max_{\alpha}\left|\mu_A(\alpha)-1/3\right| & \text{otherwise}\end{cases}$$

- $\phi(v_A,v_B) = 0$
- μ_{A1} = apply(v_A,choose) = {(y=0 z=⊥,1/3 –ε), (y=1 z=⊥,1/3 –ε), (y=2 z=⊥,1/3 +2ε)}
- μ_{B1} = apply(v_B,⊥) = v_B = {(z=⊥,1)}
- $\phi(\mu_{A1}, \mu_{B1})$ = max{ε,2ε} = 2ε
- μ_{A2} = apply(μ_{A1} ,comp) = {(y=0 z=1,1/3 –ε), (y=1 z=2,1/3 –ε), (y=2 z=0,1/3 +2ε)}
- μ_{B2} = apply(μ_{B1} ,comp) = {(z=0,1/3), (z=1,1/3), (z=2,1/3)}
- $\phi(\mu_{A2}, \mu_{B2})$ = 2/3 as the z's are different

Will return to this

Expanded Approximate Simulation Relation (EA)

Expansion of $\phi(X,Y) \rightarrow \Re_{\geq 0} \cup \{\infty\}$ is

$$\widehat{\phi}(x_1, y_1) = \min_{\substack{\psi \in D(X \times Y) \\ x_1 = \mathbf{E}[\psi_x] \\ y_1 = \mathbf{E}[\psi_y]}} \left[\max_{(x,y) \in \mathrm{supp}(\psi)} \phi(x,y) \right]$$

Expansion of:

$R \subseteq X \times Y$

(x_1, y_1) are in E(R) if there exists w, n_1, n_2

1. $w(x_1,y_1) > 0 \Rightarrow x_1 \; R \; y_1$
2. $\Sigma_y \, w(x,y) = n_1(x)$
3. $\Sigma_x \, n_1(x) = x_1$

Expanded Approximate Simulation Relation (EA)

Expansion of $\phi(X,Y) \rightarrow \Re_{\geq 0} \cup \{\infty\}$ is

$$\widehat{\phi}(x_1, y_1) = \min_{\substack{\psi \in D(X \times Y) \\ x_1 = \mathbf{E}[\psi_x] \\ y_1 = \mathbf{E}[\psi_y]}} \left[\max_{(\mathrm{x,y}) \in \mathrm{supp}(\psi)} \phi(x, y) \right]$$

$\hat{\phi}(x_1, y_1) \leq \varepsilon \Leftrightarrow \exists$ a witnessing joint distribution $\psi \in D(X \times Y)$

such that $\max\limits_{\mathrm{x,y} \in \mathrm{supp}(\psi)} \phi(x, y) \leq \varepsilon$

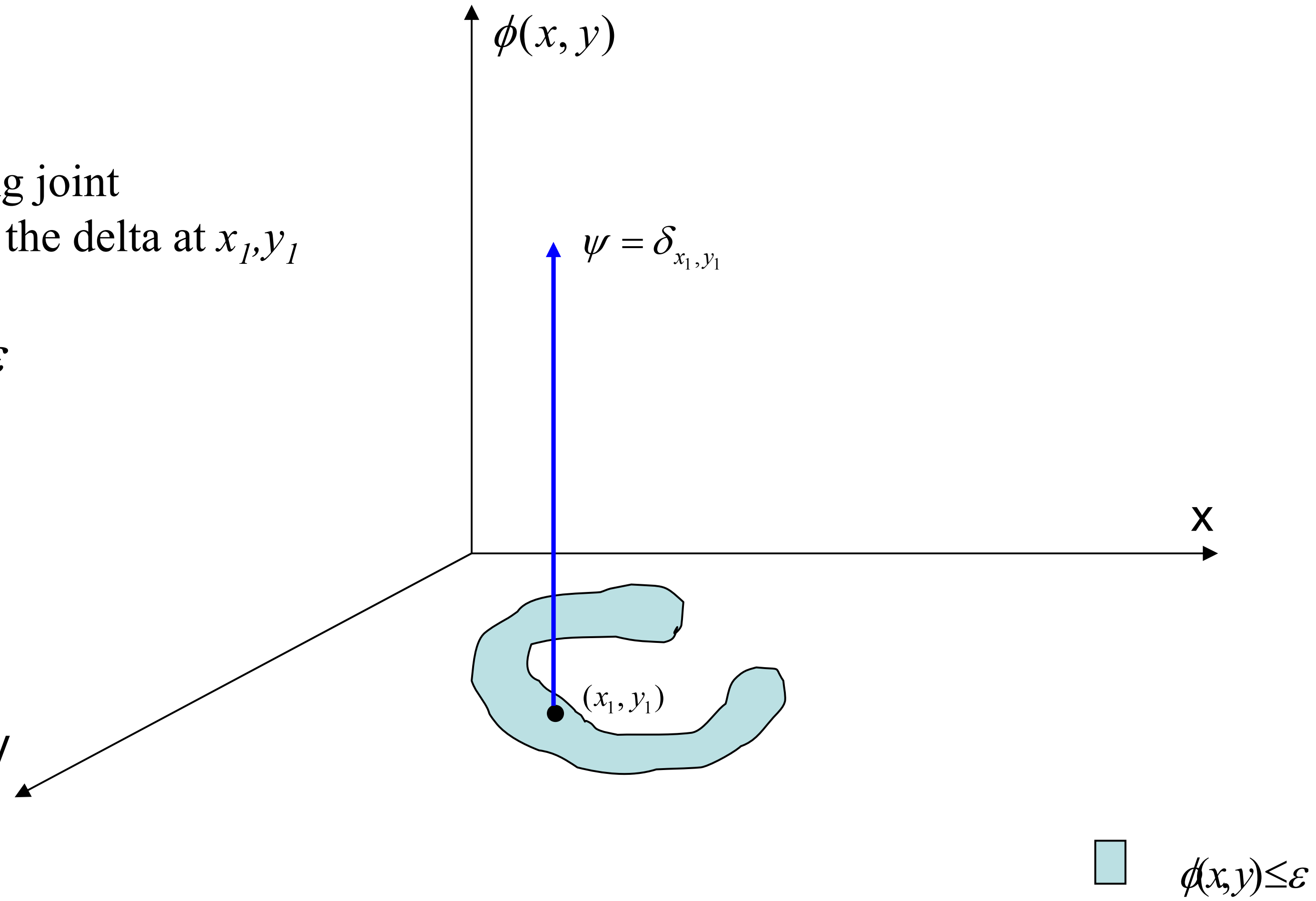
$x_1 = \sum\limits_{x,y} \psi(x, y) x$ and $y_1 = \sum\limits_{x,y} \psi(x, y) y$

Expansion

$$\phi(x_1, y_1) \leq \varepsilon$$

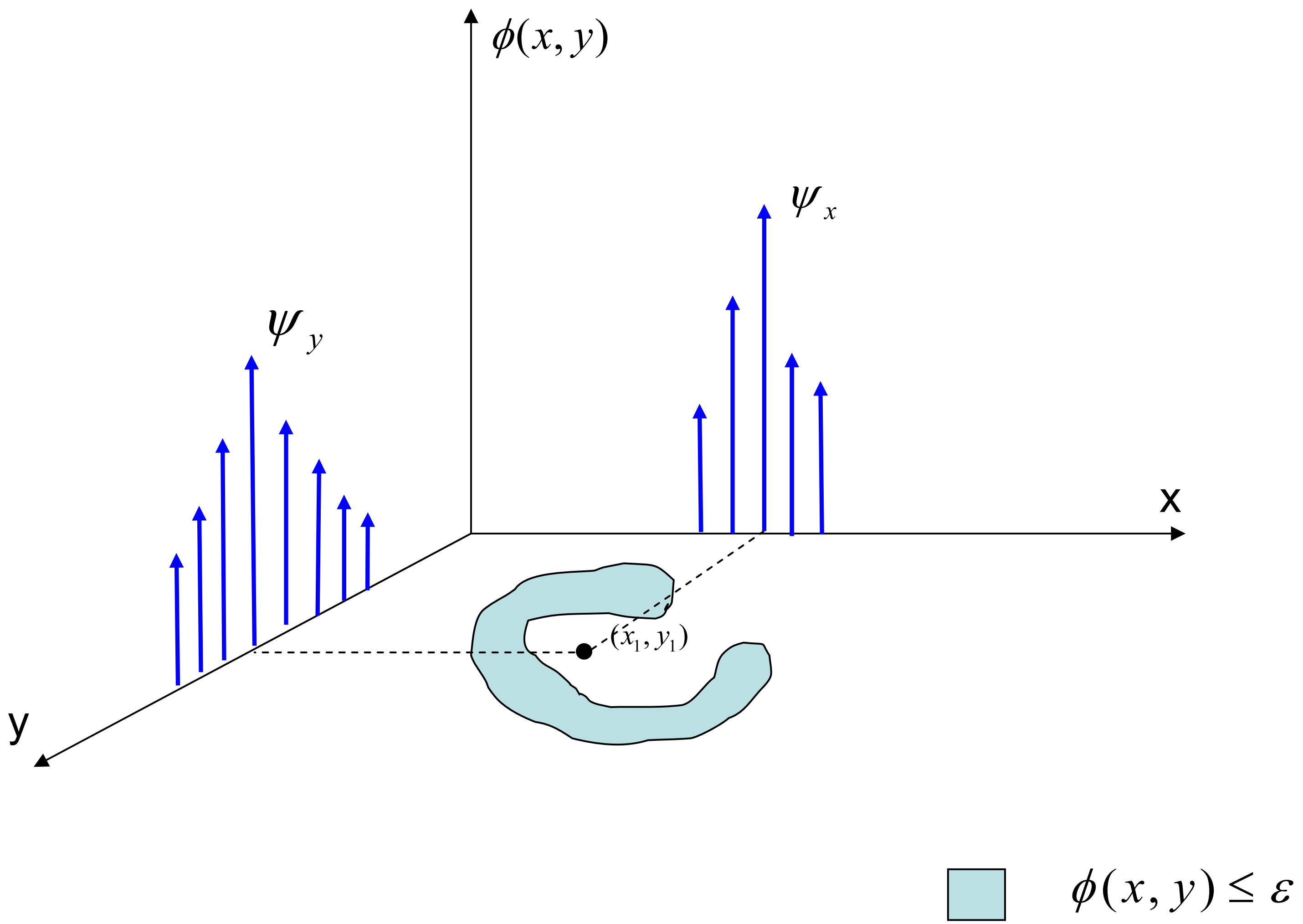
One witnessing joint distribution is the delta at x_1, y_1

$$\hat{\phi}(x_1, y_1) \leq \varepsilon$$



$$\phi(x, y) > \varepsilon$$

$$\frown{\phi}(x, y) \leq \varepsilon$$



A small digression

- **Finding the witness is an LP**
- **Optimal transportation problem (Kantorovich 1942)**

Let M be the set of all probability distributions over executions of A &B

Given $\phi(\eta, \nu)$, the pth Wasserstein metric is given by

$$w_p(\mu_1, \mu_2) = \left(\inf_{\psi \in \Gamma(\mu_1, \mu_2)} \int_{M \times M} \phi(\eta, \nu)^p \, d\psi(\eta, \nu) \right)^{\frac{1}{p}}$$

$$w_\infty(\mu_1, \mu_2) = \left(\inf_{\psi \in \Gamma(\mu_1, \mu_2)} \sup_{\eta, \nu \in \mathrm{supp}(\psi)} \phi(\eta, \nu) \right)$$

Expanded Approximate Simulation Relation (EA)

Given $\varepsilon, \delta > 0$, a function $\phi : Disc(Exec^*(A)) \times Disc(Exec^*(B)) \rightarrow \Re_{\geq 0} \cup \{\infty\}$

is a Expanded Approximate (EA) Simulation function if

1. Start : $\phi(\nu_1, \nu_2) \leq \varepsilon$

2. Step : if $\phi(\mu_1, \mu_2) \leq \varepsilon$ then $\hat{\phi}(\mu_1', \mu_2') \leq \varepsilon$

3. Trace : $\phi(\mu_1, \mu_2) \leq \varepsilon$ implies $\mathbf{d_u}(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta$

Expanded Approximate Simulation Relation (EA)

Given $\varepsilon, \delta > 0$, a function $\phi : Disc(Exec^*(A)) \times Disc(Exec^*(B)) \rightarrow \Re_{\geq 0} \cup \{\infty\}$

is a Expanded Approximate (EA) Simulation function if

1. Start : $\phi(\nu_1, \nu_2) \leq \varepsilon$

2. Step : There exists a function $c : R_1^* \times R_1 \rightarrow R_2^*$ such that,

if $\phi(\mu_1, \mu_2) \leq \varepsilon$,

σ is a task schedule and T is a task for A,

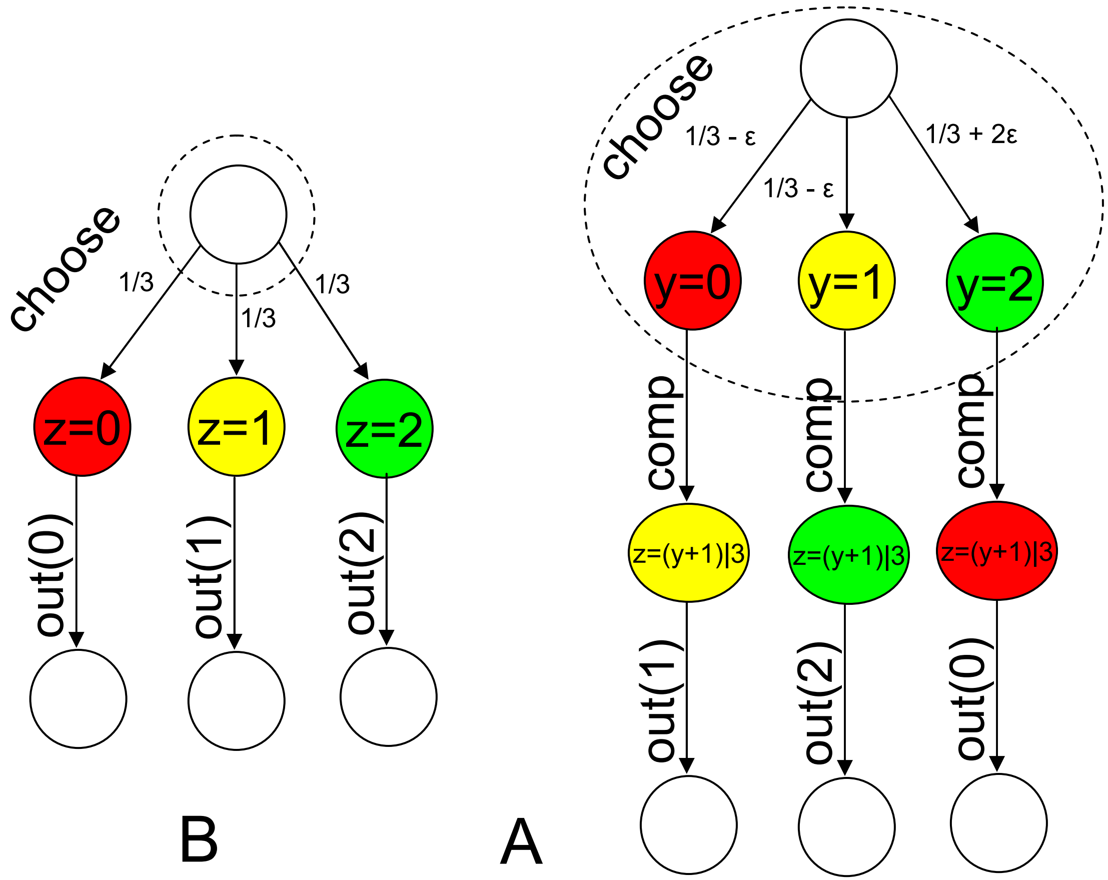
and μ_1 is consistent with σ

and μ_2 is consistent with $full(c)(\sigma)$

then $\hat{\phi}(apply(\mu_1, T), apply(\mu_2, c(\sigma, T))) \leq \varepsilon$

3. Trace : $\phi(\mu_1, \mu_2) \leq \varepsilon$ implies $\mathbf{d_u}(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta$

Return to Example



s, u : support of lsate(μ_A), lstate(μ_B)

$$\phi(\mu_A, \mu_B) = \begin{cases} \max_{\alpha,\beta} \mu_A(\alpha) + \mu_B(\beta) & \text{if } \exists \text{s}, \text{u}, \ \text{s.z} \neq \text{u.z} \\ 0 & \text{if } \forall \text{s}, \text{u}, \ \text{s.z} = \text{u.z} = \bot \text{ and s.y} = \bot \\ \max_{\alpha} \left| \mu_A(\alpha) - 1/3 \right| & \text{otherwise} \end{cases}$$

$$\phi(\mu_A,\mu_B)=\begin{cases}\max_{\alpha,\beta}\mu_A(\alpha)+\mu_B(\beta) & \text{if } \exists \text{s,u, s.z} \neq \text{u.z}\\ 0 & \text{if } \forall \text{s,u, s.z} = \text{u.z} = \bot \text{ and s.y} = \bot \\ \max_{\alpha}\left|\mu_A(\alpha)-1/3\right| & \text{otherwise}\end{cases}$$

- $\phi(v_A,v_B) = 0$
- μ_{A1} = apply(v_A,choose) = {(y=0 z=⊥,1/3 –ε), (y=1 z=⊥,1/3 –ε), (y=2 z=⊥,1/3 +2ε)}
- μ_{B1} = apply(v_B,⊥) = v_B = {z=⊥,1}
- $\phi(\mu_{A1}, \mu_{B1})$ = max{ε,2ε} = 2ε
- μ_{A2} = apply(μ_{A1} ,comp) = {(y=0 z=1,1/3 –ε), (y=1 z=2,1/3 –ε), (y=2 z=0,1/3 +2ε)}
- μ_{B2} = apply(μ_{B1} ,comp) = {(z=0,1/3), (z=1,1/3), (z=2,1/3)}
- $\phi(\mu_{A2}, \mu_{B2})$ = 2/3 as the z's are different
- For each $\rho_1 \rho_2$ in the support of ψ, $\phi(\rho_1, \rho_2) \leq 2\varepsilon$
- $\mu_{A2} = \Sigma_i \, \eta_1 \, \delta_{i, (i+1)|3}$
- $\mu_{B2} = \Sigma_i \, \eta_2 \, \delta_{i,}$

$$\hat{\phi}(\mu_{A2},\mu_{B2}) \leq 2\varepsilon$$

ψ

	$\delta_{y=0z=1}$	$\delta_{y=1z=2}$	$\delta_{y=2z=0}$
$\delta_{z=0}$	0	0	1/3
$\delta_{z=1}$	1/3 - ε	0	ε
$\delta_{z=2}$	0	1/3 - ε	ε

Key Lemmas

Lemma 1. $\widehat{\phi}(\mu_1, \mu_2) \leq \varepsilon$ with witness ψ. f_1, f_2 are distributive functions.
If $\forall \rho_1, \rho_2 \in \text{supp}(\psi), \widehat{\phi}(f_1(\rho_1), f_2(\rho_2)) \leq \varepsilon$ then
$\widehat{\phi}(f_1(\mu_1), f_2(\mu_2)) \leq \varepsilon$.

Lemma 2. $\widehat{\phi}(\mu_1, \mu_2) \leq \varepsilon$ implies $\mathbf{d_u}(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta$

$\widehat{\phi}(\mu_1, \mu_2) \leq \varepsilon$ with witness ψ. f_1, f_2 are distributive functions.

If $\forall \rho_1, \rho_2 \in \mathrm{supp}(\psi), \widehat{\phi}(f_1(\rho_1), f_2(\rho_2)) \leq \varepsilon$ then

$\widehat{\phi}(f_1(\mu_1), f_2(\mu_2)) \leq \varepsilon.$

sketch of proof for Lemma 1

For each $\rho_1, \rho_2 \in \mathrm{supp}(\psi)$, let $\psi_{\rho_1, \rho}$ be the witnessing joint for $\widehat{\phi}(f_1(\rho_1), f_2(\rho_2)) \leq \varepsilon$.

Define a new joint distribution $\psi' := \sum_{\rho_1, \rho_2 \in \mathrm{supp}(\psi)} \psi(\rho_1, \rho_2) \psi_{\rho_1, \rho_2}$

Show : $f_i(\mu_i) = \sum_{\eta_1, \eta_2} \psi'(\eta_1, \eta_2) \eta_{\mathrm{i}}$

and $\eta_1, \eta_2 \in \mathrm{supp}(\psi')$ implies $\phi(\eta_1, \eta_2) \leq \varepsilon$

$$\widehat{\phi}(\mu_1, \mu_2) \le \varepsilon \text{ implies } \mathbf{d_u}(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \le \delta$$

sketch of proof for Lemma 2

Let ψ be the witness

$$\forall \eta_1, \eta_2 \;\; \psi(\eta_1, \eta_2) \le \varepsilon$$

$$\mu_1 = \sum_{\eta_1, \eta_2} \psi(\eta_1, \eta_2)\eta_1$$

$$tdist(\mu_1) = \sum_{\eta_1, \eta_2} \psi(\eta_1, \eta_2) tdist(\eta_1)$$

$$\mathrm{d_u}(tdist(\mu_1), tdist(\mu_2)) = \operatorname*{Sup}_{\beta \in \mathrm{E}^*} \left[\sum_{\eta_1, \eta_2} \psi(\eta_1, \eta_2) tdist(\eta_1) - \sum_{\eta_1, \eta_2} \psi(\eta_1, \eta_2) tdist(\eta_2) \right]$$

$$\le \operatorname*{Sup}_{\beta \in \mathrm{E}^*} \sum_{\eta_1, \eta_2} \psi(\eta_1, \eta_2) \left| tdist(\eta_1) - tdist(\eta_2) \right| \le \delta$$

Soundness of SA

Theorem 1. If there exists an (ε,δ)-SA simulation function from A_1 to A_2 then $A_1 \leq_\delta A_2$.

- Consider any $\mu_1 = \text{apply}(v_1, T_1 T_2 \ldots T_n)$
- Define:

  $\sigma_j = c(T_1 \ldots T_j)$

  $\mu_{1,j} = \text{apply}(v_1, T_1 T_2 \ldots T_j)$

  $\mu_{2,j} = \text{apply}(v_1, \sigma_1 \sigma_2 \ldots \sigma_j)$

  $\mu_2 = \mu_{2,n}$
- For all j, $\phi(\mu_{1,j}, \mu_{2,j}) \leq \varepsilon$ (by induction)
- For all j, $d_u(\text{tdist}(\mu_{1,j}), \text{tdist}(\mu_{2,j})) \leq \delta$ (by 3.)
  - In particular, $d_u(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta$

Soundness of EA

Theorem 2. If there exists an (ε,δ)-EA simulation function from A_1 to A_2 then $A_1 \leq_\delta A_2$.

- Consider any μ_1 = apply(v_1,$T_1 T_2 \ldots T_n$)
- Define: $\sigma_j = c(T_1 \ldots T_j)$, $\mu_{1,j}$ = apply(v_1,$T_1 T_2 \ldots T_j$), $\mu_{2,j}$ = apply(v_1,$\sigma_1 \sigma_2 \ldots \sigma_j$), $\mu_2 = \mu_{2,n}$
- For all j, $\phi(\mu_{1,j},\mu_{2,j}) \leq \varepsilon$ (by Lemma 1)

$$\widehat{\phi}(\mu_{1,0},\mu_{2,0}) = \widehat{\phi}(v_1,v_2) \leq \varepsilon \text{ by start condition}$$

$$\left.\begin{aligned} \mu_{1,j+1} &= \text{apply}(\mu_{1,j},T_j) \\ \mu_{2,j+1} &= \text{apply}(\mu_{2,j},\sigma_j) \end{aligned}\right\} \text{distributive functions}$$

$$\phi(\mu_{1,j},\mu_{2,j}) \leq \varepsilon \text{ implies that } \hat{\phi}(\mu_{1,j+1},\mu_{2,j+1}) \leq \varepsilon \text{ by step condition}$$

- For all j, d_u(tdist($\mu_{1,j}$),tdist($\mu_{2,j}$)) ≤ δ (by Lemma 2)
  - In particular, d_u(tdist(μ_1),tdist(μ_2)) ≤ δ

Need for Discounting

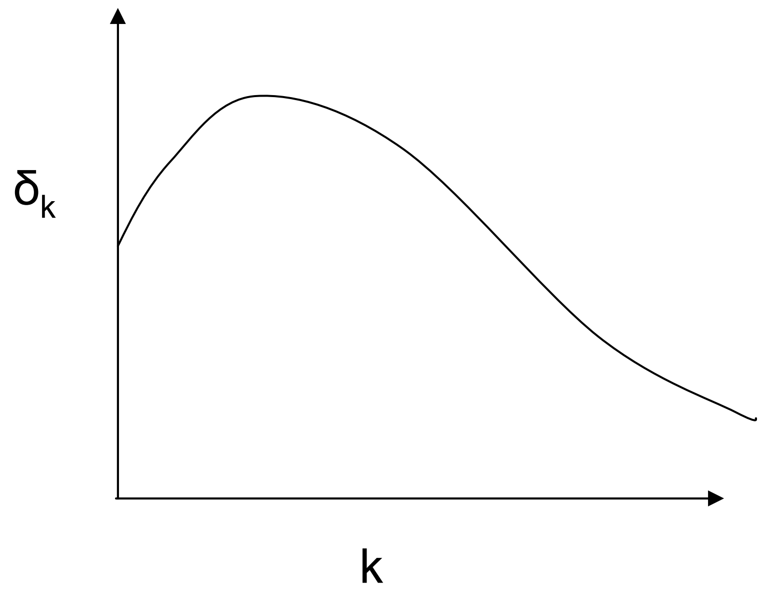
- The error $|\mu_1(\beta)-\mu_2(\beta)|$ can be small because $\mu_1(\beta)$ and $\mu_2(\beta)$ are small
- Typically when β is long
- Discounted uniform metric:

$$\mathbf{d}_k(\mu_1,\mu_2) = \sup_{\beta \in E^*, |\beta|=k} |\mu_1(\beta)-\mu_2(\beta)|$$

Given $\{\delta_k\}$ A δ_k-*implements* B if for every trace dist μ_1 of A there is a trace dist μ_2 of B such that $\mathbf{d}_k(\mu_1,\mu_2) \leq \delta_k$.

Write this as $A_1 \leq_{\delta k} A_2$.

Discount Factors



Discounted Approximate Simulation Relation (DA)

Given $\varepsilon_k, \delta_k > 0$, a collection $\{\phi_k\}, \phi_k : Disc(Exec^*(A)) \times Disc(Exec^*(B)) \to \Re_{\geq 0} \cup \{\infty\}$

is a Discounted Approximate (DA) Simulation if

1. Start : $\phi_0(\nu_1, \nu_2) \leq \varepsilon_0$

2. Step : if for all $k \leq L(\mu_1, \mu_2)$, $\phi_k(\mu_1, \mu_2) \leq \varepsilon_k$ then

then for all $k \leq L(\mu_1', \mu_2')$, $\phi_k(\mu_1', \mu_2') \leq \varepsilon_k$

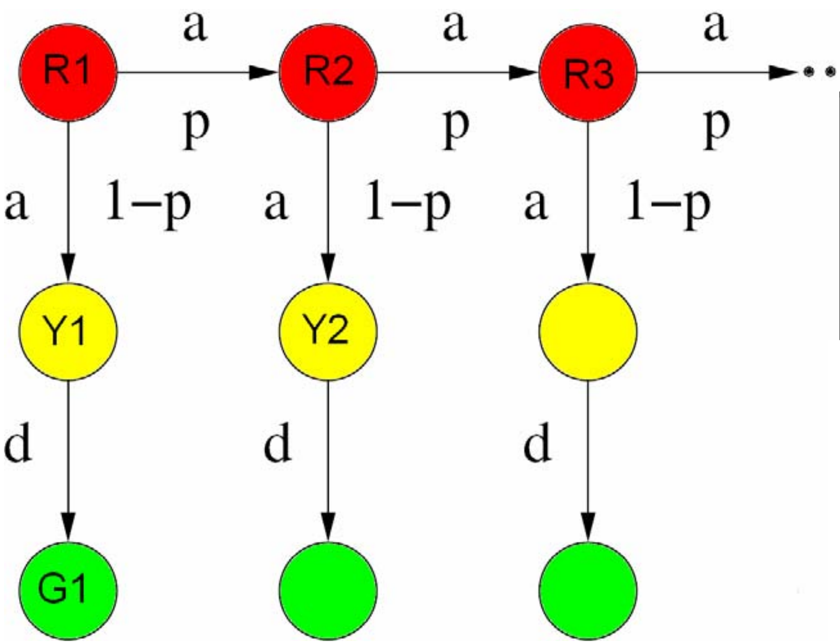
3. Trace : for all $k \leq L(\mu_1, \mu_2)$, $\phi_k(\mu_1, \mu_2) \leq \varepsilon_k$ implies

for all $k \leq L(\text{tdist}(\mu_1), \text{tdist}(\mu_2))$, $\mathbf{d}_k(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta_k$

Soundness of DA

Theorem 2. If there exists an (ε_k,δ_k)-DA simulation functions from A_1 to A_2 then $A_1 \leq_{\delta k} A_2$.

Round-based Randomized Consensus



α / σ										
⊥	1									
a		1-p	p							

$p_1(\alpha) = \mu(\alpha')\eta(q)$ if $\alpha = \alpha' a q$ and $a \in T$ and $\text{lstate}(\alpha') \rightarrow_a \eta$

$p_2(\alpha) = \mu(\alpha)$ if T is *not* enabled in $\text{lstate}(\alpha)$

Example

- A: protocol with unbiased coins
- B: protocol biased coins; p→ p+e, (1-p)→(1-p-e)
- $\phi_k(\mu_A,\mu_B) = \max_{|\alpha|=k} |\mu_A(\alpha)-\mu_B(\alpha)|$
- $\varepsilon_k = (p+e)^k - p^k$
- $\delta_k = \varepsilon_k$
- $A =_{\delta k} B$

- Step condition

  $\mu_A(\alpha)(p+e) - \mu_B(\alpha)p = p(\mu_A(\alpha) - \mu_B(\alpha)) + e\mu_A(\alpha)$

  $\leq p((p+e)^k - p^k) + e(p+e)^k$

  $= (p+e)^{k+1} - p^{k+1} = \varepsilon_k$

- Trace condition

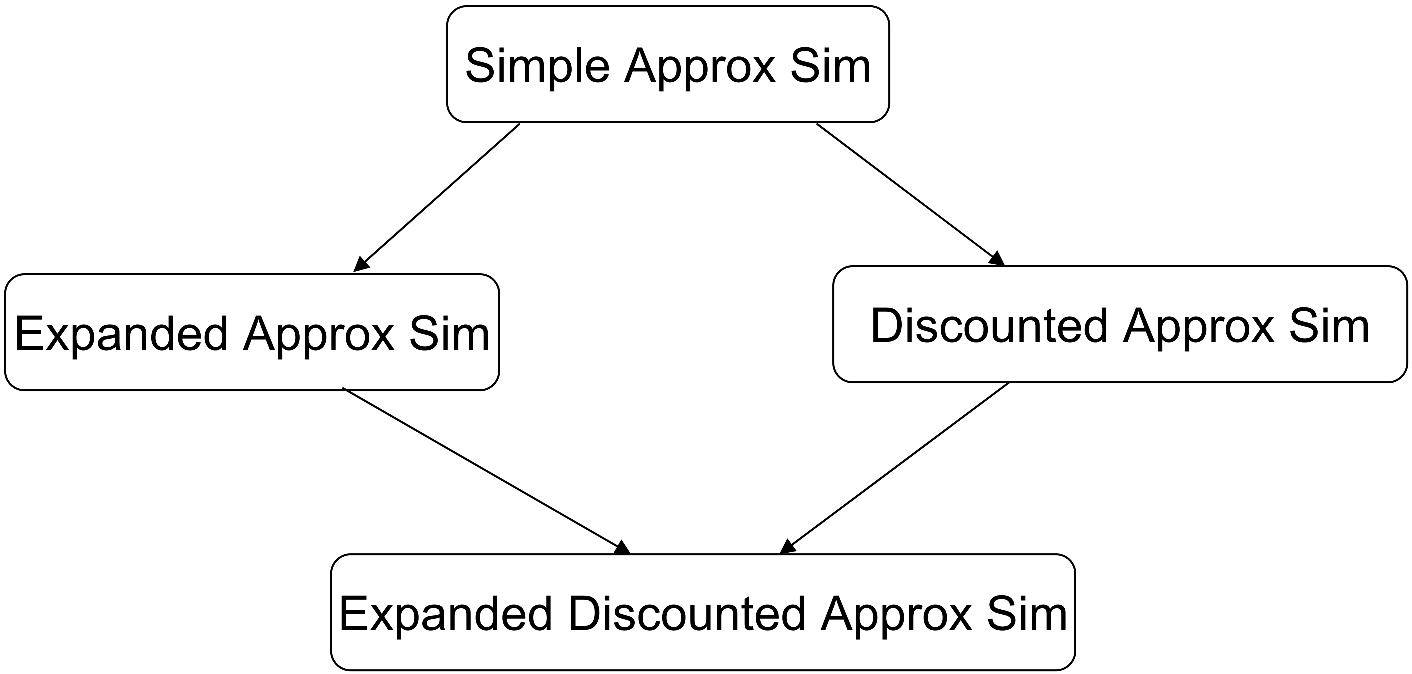
  If trace(α_1) = trace(α_2) = β then

  $|\mathrm{tdist}(\mu_A)(\beta) - \mathrm{tdist}(\mu_B)(\beta)|$

  $= |\mu_A(\alpha_1) + \mu_A(\alpha_2) - \mu_B(\alpha_1) - \mu_B(\alpha_2)|$

  $= |\mu_A(\alpha') - \mu_B(\alpha')|$, where α is the common prefix of α_1 α_2

Expanded and Discounted



Generalization

- Approximate simulations for Probabilistic Times I/O Automata
  - Q➔ X
  - Disc(X) ➔ P(X)
  - Trajectories
- Expansion:
  - $\mu(\alpha) = \Sigma_i\, \eta(\mu_i)\, \mu_i\, (\alpha)$, where η is Disc(Disc(Exec*))
  - $\mu(\alpha) = \int \mu^*(\alpha)\, d(\eta(\mu^*))$, where η is P(P(Exec*))

Metrics

Let M be the set of all probability distributions over executions of A &B

Given $\phi(\eta,\nu)$, the pth Wasserstein metric is given by

$$w_p(\mu_1,\mu_2) = \left(\inf_{\psi \in \Gamma(\mu_1,\mu_2)} \int_{M \times M} \phi(\eta,\nu)^p \, d\psi(\eta,\nu) \right)^{\frac{1}{p}}$$

$$w_\infty(\mu_1,\mu_2) = \left(\inf_{\psi \in \Gamma(\mu_1,\mu_2)} \sup_{\eta,\nu \in \mathrm{supp}(\psi)} \phi(\eta,\nu) \right)$$

Duality

$$w_1(\mu_1,\mu_2) = \sup_{f:M \to [-1,1]} \int_M f(\beta) d(\mu_1(\beta) - \mu_2(\beta))$$

Acknowledgement

- Dilsun Kaynar
- Sanjoy Mitter