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) ⊆ 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 $X \times Y$
• Forward simulation

```
\[ \text{Trace}(\alpha) = a \]
```
Fragility of Classical Implementation in Timed Systems

- Traces(A) = \{ [0,t],[0,5]a,…\}
- Traces(B) = \{ [0,t],[0,5+\varepsilon]a,…,[0,t]b,…\}
- A and B cannot be compared
and in Probabilistic Systems

- Suppose $(a_1, p_1)$ and $(a_2, p_2)$ are the only traces of $A_1$ and $A_2$
- Traces(A) contains $(a_1, rp_1) (a_2, (1-r)p_2)$
- Traces(B) contains $(a_1, (r+\epsilon)p_1) (a_2, (1-r-\epsilon)p_2)$
- Again, $A \neq B$ cannot be compared
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 $\mu_A$ in Trace($A$) there is $\mu_B$ 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 $\delta$ is unimportant
- Abstraction and model reduction through approximation

- Continuity of quantitative properties
• Define metrics for trace distributions of PIOA
• Simulation based techniques for proving $A \delta$-implements $B$

Outline
• Task-structured PIOA
• A simple approximate simulation
• Expanded approx. sim.
• Discounted approx. sim.
• Other metrics and new directions
1. Probabilistic CCS  
   Giacalone, Jou, Smolka (1994)  
   Introduced metric in the study of bisimilarity  
   Metric: $\varepsilon$-bisimilar

2. Probabilistic Concurrent 2-player Games  
   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

   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
   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 \([0,1]^{\text{Traces}^*},d\)
   Finite tests are sufficient to distinguish infinite processes
Task-Structured PTIOA
Canetti, Cheung, Kaynar, Liskov, Lynch, Pereira, Segala(2005-06)

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

Disc($X$), Supp($\mu$)

$A = (Q, v, I, O, H, D, R)$
- $Q$: countable set of states
- $v$: initial distribution
- $I, O, H$: countable, pairwise disjoint sets of actions
- $D \subseteq Q \times A \times Disc(Q)$, $(q, a, \mu) \in D$ 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 $a \in T$ enabled at $q$.

For this talk assume $A$ is closed
Executions and Traces

As usual

• Execution fragment $\alpha = q_0a_1q_1a_2\ldots$
• $\alpha$ is an execution if $q_0$ in $\text{Supp}(v)$
• Execs, Execs*
• trace($\alpha$) captures the visible part of $\alpha$
  – 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 $T_1, T_2, \ldots$
  - 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
- $\sigma$-field on Execs generated by cones
- discrete $\sigma$-field of Execs* is contained in the above
- Likewise $\sigma$-fields for Traces
- Construct chains of measures and then take limits
Applying a Schedule

Given a distribution \( \mu \) over \( \text{Execs}^* \) a task schedule \( \sigma = T_1 T_2 \ldots T_n \)
\( \text{apply}(\mu, \sigma) \) gives a probability distribution \( \text{Execs} \) by applying \( \sigma \)

- \( \text{apply}(\mu, \bot) = \mu \)
- \( \text{apply}(\mu, T) = \mu' \)
- \( \text{apply}(\mu, \sigma T) = \text{apply}(\text{apply}(\mu, \sigma), T) \)
- \( \text{apply}(\mu, \sigma) = \lim_{i \to \infty} \text{apply}(\mu, \sigma_i), \sigma_i \) is the length \( i \) prefix of \( \sigma \)

Canetti, Cheung, Kaynar, Liskov, Lynch, Pereira, Segala(05, 06)
• apply(µ, σ) = probability distribution over fragments
  
  – apply(µ, ⊥) = µ

  – apply(µ, T) = µ’

  • µ’(α) = p₁(α) + p₂(α)
  • p₁(α) = µ(α’)η(q) if α = α’aq and a is in task T and lstate(α’) ⇒ₐ η
  • p₂(α) = µ(α) if T is not enabled in lstate(α)

  – apply(µ, σT) = apply(apply(µ, σ), T )
Applying Schedules

\[ p_1(\alpha) = \mu(\alpha') p \]
\[ = \mu(\alpha')(1-p) \]
\[ = \mu(\alpha) \]

- if \( \alpha = \alpha' \) a \( R_{i+1} \) \( \text{Istate}(\alpha')=R_i \)
- if \( \alpha = \alpha' \) a \( Y_i \) \( \text{Istate}(\alpha')=R_i \)
- if \( \alpha = \alpha' \) d \( G_i \) \( \text{Istate}(\alpha')=Y_i \)

\[ p_2(\alpha) = \mu(\alpha) \]
\[ = \mu(\alpha) \]

- if \( \text{Istate}(\alpha) = R_i \)
- if \( \text{Istate}(\alpha) = Y_i \)
• A task schedule $\sigma$ defines $\mu=\text{apply}(v, \sigma)$ is a *probabilistic execution*

• Corresponding *trace distribution* $\text{tdist}(\mu)(\beta) = \mu(\text{trace}^{-1}(\beta))$

• $\text{tdists}(A) = \{\text{tdist}(\text{apply}(v, \sigma)) : \sigma \text{ is a task scheduler for } A\}$
  set of all possible trace distributions

• $A \leq B$ if $\text{tdists}(A) \subseteq \text{tdists}(B)$
Metrics on Trace Distributions

A metric \( d : \text{Disc}(E^*) \times \text{Disc}(E^*) \rightarrow \mathbb{R} \)

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

We write this as \( A \leq \delta B \).

A and B are \( \delta \)-equivalent if they \( \delta \)-implement each other.

We write this as \( A = \delta B \).
Simple Approximate Simulation Relation (SA)

Given $\varepsilon, \delta > 0$, a function $\phi : Disc(\text{Exec}^*(A)) \times Disc(\text{Exec}^*(B)) \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ is a Simple Approximate (SA) Simulation function if

1. Start: $\phi(\nu_1, \nu_2) \leq \varepsilon$
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 $d(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta$

Simulation Relation: Segala (1995-96)

$R \subseteq Disc(\text{Execs}^*(A)) \times Disc(\text{Execs}^*(B))$

1. $\nu_1 R \nu_2$
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 \( \phi : \text{Disc}(\text{Exec}^*(A)) \times \text{Disc}(\text{Exec}^*(B)) \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\} \) is a Simple Approximate (SA) 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 \( \phi(\mu_1, \mu_2) \leq \varepsilon \),
   \( \sigma \) is a task schedule and \( T \) is a task for \( A \),
   and \( \mu_1 \) is consistent with \( \sigma \)
   and \( \mu_2 \) is consistent with \( \text{full}(c)(\sigma) \)
   then \( \phi(\text{apply}(\mu_1, T), \text{apply}(\mu_2, c(\sigma, T))) \leq \varepsilon \)

3. Trace : \( \phi(\mu_1, \mu_2) \leq \varepsilon \) implies \( d_u(tdist(\mu_1), tdist(\mu_2)) \leq \delta \)
Soundness of SA

**Theorem 1.** If there exists an \((\varepsilon, \delta)\)-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.

\[ \lim_{j \to \infty} \eta_{A_j} = \eta_A \text{ and } \lim_{j \to \infty} \eta_{B_j} = \eta_B \text{ then } \lim_{j \to \infty} d(\eta_{A_j}, \eta_{B_j}) = d(\eta_A, \eta_B) \]

• **Step** and **Trace** conditions are critical for the choice of the metric and the simulation function
**Probabilistic Safety**

\[ d_u(\mu_1, \mu_2) = \sup_{C \subseteq \text{Traces}^*} |\mu_1(C) - \mu_2(C)| \]

Define a function (random variable) \( X : \text{Traces} \rightarrow \{0, 1\} \)

\( X(\beta) := 1 \) if some bad action occurs in \( \beta \)

\( 0 \) otherwise

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

Claim: \( B \) is safe with probability at least \( \delta + p \)

Let \( \mu_B \) be any trace distribution of \( B \).
There exists \( \mu_A \) such that, for all \( C \), \(|\mu_B(C) - \mu_A(C)| \leq \delta \).
Then, \( \mu_B([X=1]) \leq \delta + \mu_A([X=1]) \leq \delta + p \)
SA Simulation

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

- So, decompose $\mu_A$, $\mu_B$ into close components
  
  $\mu_A(\alpha) = \sum_i \lambda_{Ai} \mu_{Ai}(\alpha)$, where $\sum_i \lambda_{Ai} = 1$
  
  $\mu_B(\alpha) = \sum_i \lambda_{Bi} \mu_{Bi}(\alpha)$, where $\sum_i \lambda_{Bi} = 1$

  $\phi(\mu_{Ai}, \mu_{Bj})$ 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 $\phi$ gives us a measure of how close two distributions are
• We want to capture internal branching

• So, decompose $\mu_A, \mu_B$ into close components

  - $\mu_A(\alpha) = \sum_i \eta_A(\mu_{Ai}) \mu_{Ai}(\alpha)$, where $\eta_A$ is $\text{Disc}(\text{Disc}(\text{Exec}^*_A))$
  - $\mu_B(\alpha) = \sum_i \eta_B(\mu_{Bi}) \mu_{Bi}(\alpha)$, where $\eta_B$ is $\text{Disc}(\text{Disc}(\text{Exec}^*_B))$
  - $\phi(\mu_{Ai}, \mu_{Bj})$ is small

• More powerful simulation function
Need for Expansion

\[ z = \begin{cases} 0 \quad \text{if } y = 0 \\ 1 \quad \text{if } y = 1 \\ 2 \quad \text{if } y = 2 \end{cases} \]

Developed by

\[ z = (y + 1) \times 3 \]
Need for Expansion
Need for Expansion

\[
s, u : \text{support of } \text{lstate}(\mu_A), \text{lstate}(\mu_B)
\]

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

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

Will return to this
Expanded Approximate Simulation Relation (EA)

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

$$\hat{\phi}(x_1, y_1) = \min_{\psi \in D(X \times Y)} \left[ \max_{(x,y) \in \text{supp}(\psi)} \phi(x,y) \right]$$

Expansion of:
R $\subseteq X \times Y$
$(x_1, y_1)$ are in E(R) if there exists $w, n_1, n_2$
1. $w(x_1, y_1) > 0 \Rightarrow x_1 \in R y_1$
2. $\Sigma_y w(x,y) = n_1(x)$
3. $\Sigma_x n_1(x) = x_1$
Expanded Approximate Simulation Relation (EA)

Expansion of \( \phi(X, Y) \to \mathbb{R}_{\geq 0} \cup \{\infty\} \) is

\[
\tilde{\phi}(x_1, y_1) = \min_{\psi \in D(X \times Y)} \left[ \max_{(x, y) \in \text{supp}(\psi)} \phi(x, y) \right]
\]

\[
\tilde{\phi}(x_1, y_1) \leq \varepsilon \iff \exists \text{ a witnessing joint distribution } \psi \in D(X \times Y)
\]

such that \( \max_{x, y \in \text{supp}(\psi)} \phi(x, y) \leq \varepsilon \)

\[
x_1 = \sum_{x, y} \psi(x, y)x \quad \text{and} \quad y_1 = \sum_{x, y} \psi(x, y)y
\]
Expansion

\[ \phi(x_1, y_1) \leq \varepsilon \]

One witnessing joint distribution is the delta at \( x_1, y_1 \)

\[ \tilde{\phi}(x_1, y_1) \leq \varepsilon \]
\( \phi(x, y) > \varepsilon \)

\( \tilde{\phi}(x, y) \leq \varepsilon \)
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 $p$th Wasserstein metric is given by

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

$$w_\infty(\mu_1, \mu_2) = \left( \inf_{\psi \in \Gamma(\mu_1, \mu_2)} \sup_{\eta, \nu \in \text{supp}(\psi)} \phi(\eta, \nu) \right)$$
Expanded Approximate Simulation Relation (EA)

Given $\varepsilon, \delta > 0$, a function $\phi : \text{Disc}(\text{Exec}^*(A)) \times \text{Disc}(\text{Exec}^*(B)) \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ 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 $d_u(tdist(\mu_1), tdist(\mu_2)) \leq \delta$
Expanded Approximate Simulation Relation (EA)

Given \( \varepsilon, \delta > 0 \), a function \( \phi : Disc(Exec^*(A)) \times Disc(Exec^*(B)) \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\} \) 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 \( \phi(\mu_1, \mu_2) \leq \varepsilon \),
   \( \sigma \) is a task schedule and \( T \) is a task for \( A \), and \( \mu_1 \) is consistent with \( \sigma \)
   and \( \mu_2 \) 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 \( d_u(tdist(\mu_1), tdist(\mu_2)) \leq \delta \)
Return to Example

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

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

- \( \phi(v_A, v_B) = 0 \)
- \( \mu_{A1} = \text{apply}(v_A, \text{choose}) = \{(y=0, z=\perp, 1/3 - \varepsilon), (y=1, z=\perp, 1/3 - \varepsilon), (y=2, z=\perp, 1/3 + 2\varepsilon)\} \)
- \( \mu_{B1} = \text{apply}(v_B, \perp) = v_B = \{z=\perp, 1\} \)
- \( \phi(\mu_{A1}, \mu_{B1}) = \max\{\varepsilon, 2\varepsilon\} = 2\varepsilon \)
- \( \mu_{A2} = \text{apply}(\mu_{A1}, \text{comp}) = \{(y=0, z=1, 1/3 - \varepsilon), (y=1, z=2, 1/3 - \varepsilon), (y=2, z=0, 1/3 + 2\varepsilon)\} \)
- \( \mu_{B2} = \text{apply}(\mu_{B1}, \text{comp}) = \{(z=0, 1/3), (z=1, 1/3), (z=2, 1/3)\} \)
- \( \phi(\mu_{A2}, \mu_{B2}) = 2/3 \) as the z’s are different
- For each \( \rho_1, \rho_2 \) in the support of \( \psi \), \( \phi(\rho_1, \rho_2) \leq 2\varepsilon \)

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

\[ \begin{array}{c|c|c|c}
& \delta_y=0z=1 & \delta_y=1z=2 & \delta_y=2z=0 \\
\hline
\delta_z=0 & 0 & 0 & 1/3 \\
\hline
\delta_z=1 & 1/3 - \varepsilon & 0 & \varepsilon \\
\hline
\delta_z=2 & 0 & 1/3 - \varepsilon & \varepsilon \\
\end{array} \]
Key Lemmas

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

Lemma 2. $\hat{\phi}(\mu_1, \mu_2) \leq \varepsilon$ implies $d_u(tdist(\mu_1), tdist(\mu_2)) \leq \delta$
\[ \hat{\phi}(\mu_1, \mu_2) \leq \epsilon \] with witness \( \psi \). \( f_1, f_2 \) are distributive functions.

If \( \forall \rho_1, \rho_2 \in \text{supp}(\psi), \hat{\phi}(f_1(\rho_1), f_2(\rho_2)) \leq \epsilon \) then
\[ \hat{\phi}(f_1(\mu_1), f_2(\mu_2)) \leq \epsilon. \]

**sketch of proof for Lemma 1**

For each \( \rho_1, \rho_2 \in \text{supp}(\psi) \), let \( \psi_{\rho_1, \rho} \) be the witnessing joint for \( \hat{\phi}(f_1(\rho_1), f_2(\rho_2)) \leq \epsilon \).

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

Show: \( f_i(\mu_i) = \sum_{\eta_1, \eta_2} \psi'(\eta_1, \eta_2) \eta_i \)

and \( \eta_1, \eta_2 \in \text{supp}(\psi') \) implies \( \phi(\eta_1, \eta_2) \leq \epsilon \)
\[ \tilde{\phi}(\mu_1, \mu_2) \leq \varepsilon \text{ implies } d_u(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta \]

**sketch of proof for Lemma 2**

Let \( \psi \) be the witness

\[ \forall \eta_1, \eta_2 \quad \psi(\eta_1, \eta_2) \leq \varepsilon \]

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

\[ \text{tdist}(\mu_1) = \sum_{\eta_1, \eta_2} \psi(\eta_1, \eta_2) \text{tdist}(\eta_1) \]

\[ d_u(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) = \sup_{\beta \in E^*} \left[ \sum_{\eta_1, \eta_2} \psi(\eta_1, \eta_2) \text{tdist}(\eta_1) - \sum_{\eta_1, \eta_2} \psi(\eta_1, \eta_2) \text{tdist}(\eta_2) \right] \]

\[ \leq \sup_{\beta \in E^*} \sum_{\eta_1, \eta_2} |\psi(\eta_1, \eta_2)| \text{tdist}(\eta_1) - \text{tdist}(\eta_2) | \leq \delta \]
Soundness of SA

Theorem 1. If there exists an \((\varepsilon, \delta)\)-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 \(j\), \(d_u(tdist(\mu_{1,j}), tdist(\mu_{2,j})) \leq \delta\) (by 3.)
  - In particular, \(d_u(tdist(\mu_1), tdist(\mu_2)) \leq \delta\)
Soundness of EA

**Theorem 2.** If there exists an \((\varepsilon, \delta)\)-EA 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 Lemma 1)
  \[
  \phi(\mu_{1,0}, \mu_{2,0}) = \phi(v_1, v_2) \leq \varepsilon \quad \text{by start condition}
  \]
  \[
  \mu_{1,j+1} = \text{apply}(\mu_{1,j}, T_j) \]
  \[
  \mu_{2,j+1} = \text{apply}(\mu_{2,j}, \sigma_j) \]
  distributive functions
  
  \(\phi(\mu_{1,j}, \mu_{2,j}) \leq \varepsilon\) implies that \(\phi(\mu_{1,j+1}, \mu_{2,j+1}) \leq \varepsilon\) by step condition

- For all \(j\), \(d_u(\text{tdist}(\mu_{1,j}), \text{tdist}(\mu_{2,j})) \leq \delta\) (by Lemma 2)
  - In particular, \(d_u(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta\)
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 $\beta$ is long
• Discounted uniform metric:

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

Given $\{\delta_k\}$ $A \delta_k$-implements $B$ if for every trace dist $\mu_1$ of $A$ there is a trace dist $\mu_2$ of $B$ such that $d_k(\mu_1, \mu_2) \leq \delta_k$.

Write this as $A_1 \leq_{\delta_k} A_2$. 
Discount Factors

\[ \delta_k \]

\[ k \]
Discounted Approximate Simulation Relation (DA)

Given $\varepsilon_k, \delta_k > 0$, a collection $\{\phi_k\}, \phi_k : Disc(\text{Exec}^*(A)) \times Disc(\text{Exec}^*(B)) \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ is a Discounted Approximate (DA) Simulation if

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

2. Step: if \( \text{for all } k \leq L(\mu_1, \mu_2), \phi_k(\mu_1, \mu_2) \leq \varepsilon_k \) then
   \( \text{then for all } k \leq L(\mu_1', \mu_2'), \phi_k(\mu_1', \mu_2') \leq \varepsilon_k \)

3. Trace: \( \text{for all } k \leq L(\mu_1, \mu_2), \phi_k(\mu_1, \mu_2) \leq \varepsilon_k \) implies \( \text{for all } k \leq L(\text{tdist}(\mu_1), \text{tdist}(\mu_2)), d_k(\text{tdist}(\mu_1), \text{tdist}(\mu_2)) \leq \delta_k \)
Soundness of DA

**Theorem 2.** If there exists an $(\varepsilon_k, \delta_k)$-DA simulation functions from $A_1$ to $A_2$ then $A_1 \leq_{\delta_k} A_2$. 
Round-based Randomized Consensus

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

$p_2(\alpha) = \mu(\alpha)$ if $T$ is not enabled in $lstate(\alpha)$
Example

• A: protocol with unbiased coins
• B: protocol biased coins; $p \rightarrow p+e$, $(1-p) \rightarrow (1-p-e)$
• $\phi_k(\mu_A, \mu_B) = \max_{|\alpha|=k} |\mu_A(\alpha) - \mu_B(\alpha)|$
• $\varepsilon_k = (p+e)^k - p^k$
• $\delta_k = \varepsilon_k$
• $A = \delta_k B$
• 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 \( \text{trace}(\alpha_1) = \text{trace}(\alpha_2) = \beta \) then
\[ |\text{tdist}(\mu_A)(\beta) - \text{tdist}(\mu_B)(\beta)| \]
\[ = |\mu_A(\alpha_1) + \mu_A(\alpha_2) - \mu_B(\alpha_1) - \mu_B(\alpha_2)| \]
\[ = |\mu_A(\alpha') - \mu_B(\alpha')|, \text{ where } \alpha \text{ is the common prefix of } \alpha_1 \alpha_2 \]
Expanded and Discounted

- Simple Approx Sim
- Expanded Approx Sim
- Discounted Approx Sim
- Expanded Discounted Approx Sim
Generalization

- Approximate simulations for Probabilistic Times I/O Automata
  - $Q \Rightarrow X$
  - $\text{Disc}(X) \Rightarrow P(X)$
  - Trajectories

- Expansion:
  - $\mu(\alpha) = \sum_i \eta(\mu_i) \mu_i(\alpha)$, where $\eta$ is $\text{Disc}(\text{Disc}(\text{Exec}*))$
  - $\mu(\alpha) = \int \mu^*(\alpha) \, d(\eta(\mu*))$, where $\eta$ is $P(P(\text{Exec}*))$
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

\[
w_p(\mu_1, \mu_2) = \left( \inf_{\psi \in \Gamma(\mu_1, \mu_2)} \int \phi(\eta, \nu)^p \, d\psi(\eta, \nu) \right)^{\frac{1}{p}}
\]

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

Duality

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

• Dilsun Kaynar
• Sanjoy Mitter