Approximate Testing Equivalence Based on Time, Probability, and Observed Behavior

Alessandro Aldini
University of Urbino "Carlo Bo", Italy

PaCo at L'Aquila, 2-3 Marzo 2010

Outline

- Why Approximate Equivalence Checking.
- Testing Semantics.
- Three views of Approximate Testing Equivalence.
- Future work.

Why Approximate Equivalence Checking

Applications of equivalence checking:

- relating a process model to a reference model;
- verifying substitutions/transformations/reductions that are expected to preserve system properties;
- noninterference analysis.

$$\neg \text{ Perfect equivalence}$$
$$\Downarrow$$
$$\text{Quantitative comparison}$$
$$\Downarrow$$
$$\text{Numbers!}$$

Most popular solution: approximating bisimulation

Why bisimulation...

- It is a relation that can be relaxed (approximate bisimulation).
- It has a suitable modal logic characterization (pseudometrics approach).

Example: Pseudometrics [Desharnais et al., vBW, ...]

- Logical characterization of bisimulation:

  $\mathcal{L} := \top \mid \phi_1 \wedge \phi_2 \mid \langle a \rangle_q \phi$

- From the logic-based characterization to the functional expressions based characterization:

  $f := \mathbf{1} \mid \mathbf{1} - f \mid \langle a \rangle f \mid \min(f_1, f_2) \mid \sup_{i \in \mathbb{N}} f_i \mid f \ominus q$

- s and s' are bisimilar iff they satisfy the same logical formulas iff they have the same values for each functional expression.

- Pseudometric: $d^c(P, Q) = sup_{f \in \mathcal{F}^c} |f_P(p_0) - f_Q(q_0)|$

Example: Pseudometrics [Desharnais et al., vBW, ...]

- Logical characterization of bisimulation:

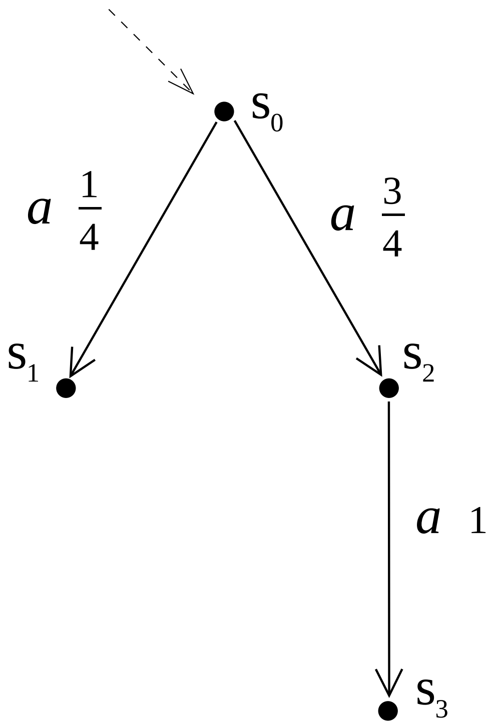
  $\mathcal{L} := \top \mid \phi_1 \wedge \phi_2 \mid \langle a \rangle_q \phi$

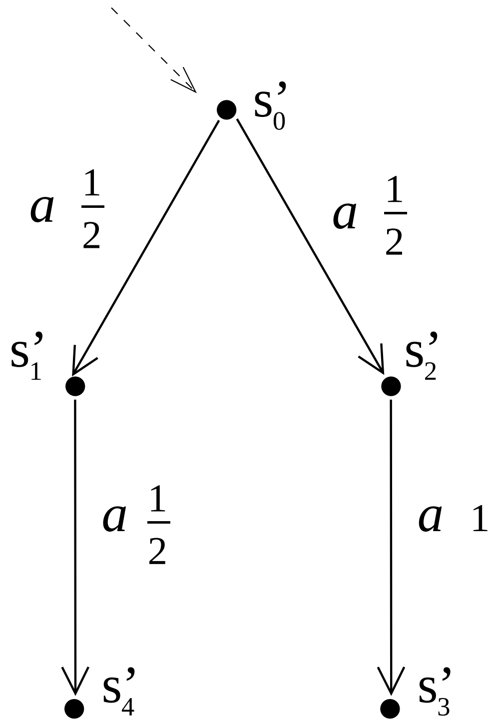
- From the logic-based characterization to the functional expressions based characterization:

  $f := \mathbf{1} \mid \mathbf{1} - f \mid \langle a \rangle f \mid \min(f_1, f_2) \mid \sup_{i \in \mathbb{N}} f_i \mid f \ominus q$

  * $\mathbf{1}(s) = 1$
  * $(\mathbf{1} - f)(s) = 1 - f(s)$
  * $\langle a \rangle f(s) = c \int_S f(t) \tau_a(s, dt)$
  * $(f \ominus q)(s) = \max(f(s) - q, 0)$

Example: Pseudometrics [Desharnais et al., vBW, ...]





$\langle a\rangle.\langle a\rangle\mathbf{1}$ evaluates to $3c^2/4$ at state s_0 and to 0 elsewhere

$\langle a\rangle.\langle a\rangle\mathbf{1}$ evaluates to $3c^2/4$ at state s_0 and to 0 elsewhere

$\langle a\rangle.(\langle a\rangle\mathbf{1} \ominus c/2)$ evaluates to $3c^2/8$ at state s_0'

$\langle a\rangle.(\langle a\rangle\mathbf{1} \ominus c/2)$ evaluates to $c^2/4$ at state s_0'

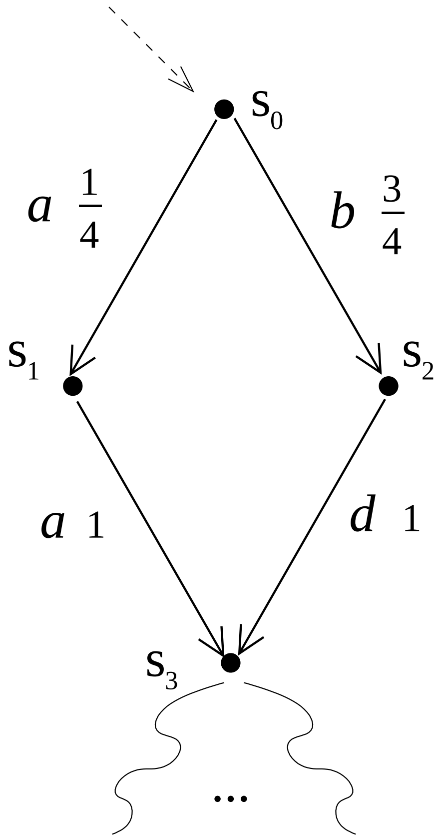
Example: Pseudometrics [Desharnais et al., vBW, ...]

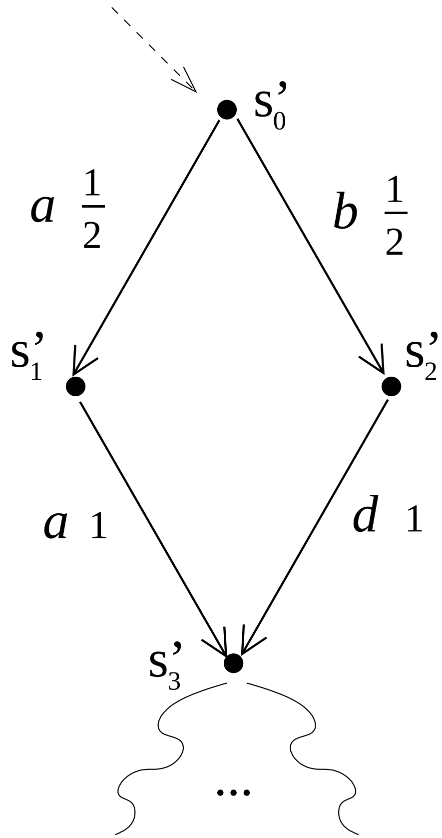
$d^c(P, Q) = sup_{f \in \mathcal{F}^c} |f_P(p_0) - f_Q(q_0)|$

Limitations concerning the interpretation of the distance:

- it is stated-based, what about an activity-oriented setting...
- any pair of states can be considered, which comparisons make sense...

Example: Pseudometrics [Desharnais et al., vBW, ...]





- if $c = 1$ then s_3 (s'_3) is as important as s_0 (s'_0)
- no functional expression reveals that the probability of reaching s_3 (s'_3) is 1

Other approaches: approx. bisimulation

A relation $R \subseteq S \times S$ is a:

1. weak probabilistic bisimulation with ε precision if whenever $(s, s') \in R$, then for all C in the partition induced by R and $\forall a \in Act.\, d(s, s', a, C) \leq \varepsilon$.
   [ADiP,Ald]

2. ε-simulation if whenever sRt, then $\forall a \in Act, X \subseteq S.\, h_a(t, R(X)) \geq h_a(s, X) - \varepsilon$. Then, R is a ε-bisimulation if it is symmetric and a ε-simulation.
   [Desh. et al.]

3. ε-bisimulation if whenever sRt, then the norm of a linear operator applied to the matrix representations of s and t with respect to a R-based classification operator is confined by ε.
   [DiPHW]

Other approaches: approx. bisimulation

1. has a clear numerical interpretation (relation with quasi-lumpability), but not a poly-time verification algorithm.

2. has logic-based and game-theoretic characterizations, a poly-time verification algorithm, but strong usability limitations.

3. is efficient, but the measure strictly depends on the chosen norms and classification linear operators.

A Different Approach

- ...based on Markovian testing equivalence.
- ...dealing with temporal and probabilistic aspects of the observed behaviors.
- ...including a quantitative comparison of the observed behaviors based on typical behaviors.

Markovian process calculus

- Actions are exp. timed: $<a, \lambda>$ with rate $\lambda \in \mathbb{R}_{>0}$ and average duration given by the inverse of the rate.
- $P ::= \underline{0} \mid <a, \lambda>.P \mid P + P \mid A$
- $\mathcal{P}$ is the set of closed and guarded process terms.
- Exit rate:

$$rate(P, a, C) \;=\; \sum \{\!| \, \lambda \in \mathbb{R}_{>0} \mid \exists P' \in C.\, P \xrightarrow{a,\lambda} P' \, |\!\}$$

$$rate_{\mathrm{t}}(P) \;=\; \sum_{a \in Name} rate(P, a, \mathcal{P})$$

Markovian process calculus: computations

Concrete trace:

$$trace(c) \;=\; \begin{cases} \delta & \text{if } |c| = 0 \\ a \circ trace(c') & \text{if } c \equiv P \xrightarrow{a,\lambda} c' \end{cases}$$

Probability:

$$prob(c) \;=\; \begin{cases} 1 & \text{if } |c| = 0 \\ \frac{\lambda}{rate_{\mathrm{t}}(P)} \cdot prob(c') & \text{if } c \equiv P \xrightarrow{a,\lambda} c' \end{cases}$$

$$prob(C) \;=\; \sum_{c \in C} prob(c)$$

Markovian process calculus: computations

Stepwise average duration:

$$time(c) = \begin{cases} \delta & \text{if } |c| = 0 \\ \frac{1}{rate_{\mathrm{t}}(P)} \circ time(c') & \text{if } c \equiv P \xrightarrow{a,\lambda} c' \end{cases}$$

Computations with stepwise average duration not greater than $\theta \in (\mathbb{R}_{>0})^*$:

$$C_{\leq \theta} = \{\!| \, c \in C \mid |c| \leq |\theta| \wedge \forall i = 1, \ldots, |c|. \, time(c)[i] \leq \theta[i] \, |\!\}.$$

C^l: computations in C whose length is equal to $l \in \mathbb{N}$.

Tests

The set $\mathbb{T}_{\mathrm{R,c}}$ of canonical reactive tests is generated by the syntax:

$$T \; ::= \; \mathrm{s} \mid {<}a, *_1{>}.T + \sum_{b \in \mathcal{E} - \{a\}} {<}b, *_1{>}.\mathrm{f}$$

where $a \in \mathcal{E}$, $\mathcal{E} \subseteq \mathit{Name} - \{\tau\}$ finite, the summation is absent whenever $\mathcal{E} = \{a\}$, and s (resp. f) is a zeroary operator standing for success (resp. failure).

- $[\![P \parallel T]\!]$, with $\parallel$ a CSP-like parallel composition operator, is called a **configuration**, which is **successful** if its test part is s.
- A test-driven computation is successful if it traverses a successful configuration.
- $\mathcal{SC}(P, T)$: multiset of successful computations of $P \parallel T$.

Markovian Testing Equivalence

Let $P_1, P_2 \in \mathcal{P}$. We say that P_1 is Markovian testing equivalent to P_2, written $P_1 \sim_{\rm MT} P_2$, iff for all reactive tests $T \in \mathbb{T}_{\rm R,c}$ and sequences $\theta \in (\mathbb{R}_{>0})^*$ of average amounts of time:

$$prob(\mathcal{SC}^{|\theta|}_{\leq \theta}(P_1, T)) \;=\; prob(\mathcal{SC}^{|\theta|}_{\leq \theta}(P_2, T)).$$

Intuition: for each test, the two sets of **observed** successful computations are characterized by the same **probabilities** and **stepwise average durations**.

Approx. Time: P_2 is a slow approx. of P_1

Intuition: the same tests are passed with the same probabilities, but the successful computations of P_2 can be slower (up to ϵ) than those of P_1.

$$C_{\leq \theta + \epsilon} \; = \; \{\!| \, c \in C \mid |c| \leq |\theta| \wedge \forall i = 1, \ldots, |c|. \, \mathit{time}(c)[i] \leq \theta[i] + \epsilon \, |\!\}.$$

Let $P_1, P_2 \in \mathcal{P}$ and $\epsilon \in \mathbb{R}_{\geq 0}$. We say that P_2 is **slow Markovian testing ϵ-similar** to P_1 iff for all reactive tests $T \in \mathbb{T}_{\mathrm{R,c}}$ and sequences $\theta \in (\mathbb{R}_{>0})^*$ of average amounts of time: $\mathit{prob}(\mathcal{SC}^{|\theta|}_{\leq \theta}(P_1, T)) \; = \; \mathit{prob}(\mathcal{SC}^{|\theta|}_{\leq \theta + \epsilon}(P_2, T))$.

- Conservative extension of $\sim_{\mathrm{MT}}$.
- "Transitive": $d(P_1, P_2) = \epsilon_1 \wedge d(P_2, P_3) = \epsilon_2 \rightarrow d(P_1, P_3) = \epsilon_1 + \epsilon_2$
- Checkable in poly-time.
- Not practical: it may happen that P_2 is s.M.t. p-similar to P_1 but not s.M.t. $(p + q)$-similar to P_1!

Approx. Time: P_2 is a slow approx. of P_1

Intuition: $\mathcal{SC}^{|\theta|}_{\leq\theta}(P_1, T)$ is compared with $\mathcal{SC}^{|\theta|}_{\leq\theta}(P_2, T)$ augmented with the successful T-driven computations of P_2 that are slower (up to ϵ) than corresponding computations in $\mathcal{SC}^{|\theta|}_{\leq\theta}(P_1, T)$.

$$\begin{aligned} C_{\leq\theta+\epsilon,C'} \;=\; & C_{\leq\theta}\cup \\ & \{\!| \, c \in C \mid c \notin C_{\leq\theta} \wedge \exists c' \in C'_{\leq\theta}.\, |c| \leq |c'| \wedge \\ & \forall i = 1, \dots, |c|.\, \mathit{time}(c')[i] \leq \mathit{time}(c)[i] \leq \mathit{time}(c')[i] + \epsilon \, |\!\}. \end{aligned}$$

Let $P_1, P_2 \in \mathcal{P}$ and $\epsilon \in \mathbb{R}_{\geq 0}$. We say that P_2 is **slow Markovian testing ϵ-similar** to P_1 iff for all reactive tests $T \in \mathbb{T}_{\mathrm{R,c}}$ and sequences $\theta \in (\mathbb{R}_{>0})^*$ of average amounts of time: $\mathit{prob}(\mathcal{SC}^{|\theta|}_{\leq\theta}(P_1, T)) \;=\; \mathit{prob}(\mathcal{SC}^{|\theta|}_{\leq\theta+\epsilon,\mathcal{SC}^{|\theta|}(P_1,T)}(P_2, T))$.

- Conservative extension of $\sim_{\mathrm{MT}}$.
- "Transitive": $d(P_1, P_2) = \epsilon_1 \wedge d(P_2, P_3) = \epsilon_2 \rightarrow d(P_1, P_3) = \delta$ with $\delta \leq \epsilon_1 + \epsilon_2$.
- Checkable in poly-time.

Example

$$<g, \gamma>.<a, \lambda>.<b, \lambda>.\underline{0} + <g, \gamma>.<a, \lambda>.<d, \lambda>.\underline{0}$$

$$<g, \gamma>.<a, \lambda>.<d, \lambda - \delta>.\underline{0} + <g, \gamma>.<a, \lambda - \delta>.<b, \lambda>.\underline{0}$$

$$\epsilon \geq \frac{1}{\lambda-\delta} - \frac{1}{\lambda}$$

Approx. Time: further definitions

- Fast approximation is obtained by a dual argument:

  Let $P_1, P_2 \in \mathcal{P}$ and $\epsilon \in \mathbb{R}_{\geq 0}$. We say that P_2 is **fast Markovian testing ϵ-similar** to P_1 iff for all reactive tests $T \in \mathbb{T}_{\mathrm{R,c}}$ and sequences $\theta \in (\mathbb{R}_{>0})^*$ of average amounts of time: $prob(\mathcal{SC}^{|\theta|}_{\leq \theta + \epsilon, \mathcal{SC}^{|\theta|}(P_2,T)}(P_1, T)) = prob(\mathcal{SC}^{|\theta|}_{\leq \theta}(P_2, T))$.

- Fast and slow approximations can be combined:

$$
\begin{aligned}
C_{\leq \theta \pm \epsilon, C'} \; = \; & C_{\leq \theta} \cup \\
& \{\!| \, c \in C \mid c \notin C_{\leq \theta} \wedge \exists c' \in C'_{\leq \theta}. \, |c| \leq |c'| \wedge \\
& \forall i = 1, \ldots, |c|. \, time(c')[i] - \epsilon \leq time(c)[i] \leq time(c')[i] + \epsilon \, |\!\}
\end{aligned}
$$

  Let $P_1, P_2 \in \mathcal{P}$ and $\epsilon \in \mathbb{R}_{\geq 0}$. We say that P_2 is **temporally Markovian testing ϵ-similar** to P_1 iff for all reactive tests $T \in \mathbb{T}_{\mathrm{R,c}}$ and sequences $\theta \in (\mathbb{R}_{>0})^*$ of average amounts of time:

$$
prob(\mathcal{SC}^{|\theta|}_{\leq \theta \pm \epsilon, \mathcal{SC}^{|\theta|}(P_2,T)}(P_1, T)) \; = \; prob(\mathcal{SC}^{|\theta|}_{\leq \theta \pm \epsilon, \mathcal{SC}^{|\theta|}(P_1,T)}(P_2, T)).
$$

Examples

$$<g,\gamma>.<a,\lambda>.<b,\lambda>.\underline{0} + <g,\gamma>.<a,\lambda>.<d,\lambda>.\underline{0}$$

$$<g,\gamma>.<a,\lambda>.<d,\lambda+\delta>.\underline{0} + <g,\gamma>.<a,\lambda+\delta>.<b,\lambda>.\underline{0}$$

$$\epsilon \geq \frac{1}{\lambda} - \frac{1}{\lambda+\delta}$$

$$<g,\gamma>.<a,\lambda>.<b,\lambda>.\underline{0} + <g,\gamma>.<a,\lambda>.<d,\lambda>.\underline{0}$$

$$<g,\gamma>.<a,\lambda-\delta>.<d,\lambda+\delta>.\underline{0} + <g,\gamma>.<a,\lambda+\delta>.<b,\lambda-\delta>.\underline{0}$$

$$\epsilon \geq \frac{1}{\lambda-\delta} - \frac{1}{\lambda}$$

Approximating Probabilities

Intuition: the same tests are passed with the same temporal constraints but with different probabilities.

Let $P_1, P_2 \in \mathcal{P}$ and $\epsilon \in \mathbb{R}_{\geq 0}$. We say that P_2 is **probabilistically Markovian testing ϵ-similar** to P_1 iff for all reactive tests $T \in \mathbb{T}_{\mathrm{R,c}}$ and sequences $\theta \in (\mathbb{R}_{>0})^*$ of average amounts of time: $|prob(\mathcal{SC}^{|\theta|}_{\leq \theta}(P_1, T)) - prob(\mathcal{SC}^{|\theta|}_{\leq \theta}(P_2, T))| \leq \epsilon$.

- This problem is undecidable.
- Relaxations of the problem can be decided (e.g. polynomially accurate similarity).

Approximating Observed Behavior

Idea: Processes are compared w.r.t. an event log describing typical behaviors and a fitness measure expressing the overlap in fitting these behaviors [de Medeiros, van der Aalst, Weijters, 2008].

Approach:

- Typical behavior → Tests satisfying a logic formula ϕ.
- Fitness measure → Similarity between tests.

Intuition: similar tests are passed with the same temporal constraints and probabilities.

Test similarity

Precision establishes whether the behavior of the second test is possible from the viewpoint of the behavior of the first test.

$$prec(T, T') \quad = \quad \frac{1}{|T'|} \sum_{i=1}^{|T'|} \frac{|(enabled(T,i,s) \cap enabled(T',i,s)) \cup (enabled(T,i,f) \cap enabled(T',i,f))|}{|enabled(T',i,f)| + |enabled(T',i,s)|}$$

Recall establishes how much of the behavior of the first test is covered by the second test.

$$rec(T, T') \quad = \quad \frac{1}{|T|} \sum_{i=1}^{|T|} \frac{|(enabled(T,i,s) \cap enabled(T',i,s)) \cup (enabled(T,i,f) \cap enabled(T',i,f))|}{|enabled(T,i,f)| + |enabled(T,i,s)|}$$

Examples

$T_1 = <a, *_1>.\mathrm{s} + <b, *_1>.\mathrm{f}$
$T_2 = <b, *_1>.\mathrm{s} + <a, *_1>.\mathrm{f}$

$$prec(T_1, T_2) = rec(T_1, T_2) = 0$$

$T_1 = <a_1, *_1>.<a_2, *_1>.\mathrm{s} + <b, *_1>.\mathrm{f}$
$T_2 = <c, *_1>.<a_2, *_1>.\mathrm{s} + <b, *_1>.\mathrm{f} + <b', *_1>.\mathrm{f}$

$$prec(T_1, T_2) = \frac{2}{3} \text{ and } rec(T_1, T_2) = \frac{3}{4}$$

Transitivity relations

$prec(T_1, T_2)$	$rec(T_1, T_2)$	$prec(T_2, T_3)$	$rec(T_2, T_3)$	$prec(T_1, T_3)$	$rec(T_1, T_3)$
z	w	x	y	≤ 1	≤ 1
z	w	x	1	< 1	$\geq w$
z	w	1	y	≤ 1	$\leq w$
z	w	1	1	z	w
z	1	x	y	$\leq x$	≤ 1
z	1	x	1	$< x$	1
z	1	1	y	≤ 1	≤ 1
z	1	1	1	z	1
1	w	x	y	$\geq x$	≤ 1
1	w	x	1	$\geq x$	$\geq w$
1	w	1	y	1	$< w$
1	w	1	1	1	w
1	1	x	y	x	y
1	1	x	1	x	1
1	1	1	y	1	y
1	1	1	1	1	1

Approx. Behavior: definitions

Attempt 1: abstracting from time...

Let $P_1, P_2 \in \mathcal{P}$ and $\mathbb{T}_{\mathrm{R,c},\phi}$ a finite set of tests. We say that P_2 is **behaviorally Markovian testing similar** to P_1 with precision $p \in [0,1]$ and recall $r \in [0,1]$ iff for each reactive test $T \in \mathbb{T}_{\mathrm{R,c},\phi}$ there exists a reactive test $T' \in \mathbb{T}_{\mathrm{R,c},\phi}$ such that:

1. $prec(T, T') \geq p$ and $rec(T, T') \geq r$
2. $prob(\mathcal{SC}(P_1, T)) = prob(\mathcal{SC}(P_2, T'))$

Attempt 2: adding time by exploiting a canonical set of average amounts of time...

Let $P_1, P_2 \in \mathcal{P}$ and $\mathbb{T}_{\mathrm{R,c},\phi}$ a finite set of tests. We say that P_2 is **behaviorally Markovian testing similar** to P_1 with precision $p \in [0,1]$ and recall $r \in [0,1]$ iff for each reactive test $T \in \mathbb{T}_{\mathrm{R,c},\phi}$ there exists a reactive test $T' \in \mathbb{T}_{\mathrm{R,c},\phi}$ such that for all sequences $\theta \in \Theta(P_1, T) \cup \Theta(P_2, T')$ of average amounts of time:

1. $prec(T, T') \geq p$ and $rec(T, T') \geq r$
2. $prob(\mathcal{SC}^{|\theta|}_{\leq \theta}(P_1, T)) = prob(\mathcal{SC}^{|\theta|}_{\leq \theta}(P_2, T'))$

Approx. Behavior: definitions

Attempt 3: relaxing all the three dimensions...

Let $P_1, P_2 \in \mathcal{P}$ and $\mathbb{T}_{\mathrm{R,c},\phi}$ a finite set of tests. We say that P_2 is **Markovian testing similar** to P_1 with precision $p \in [0,1]$, recall $r \in [0,1]$, temporal threshold $\epsilon \in \mathbb{R}_{>0}$, and probability threshold $\nu \in \mathbb{R}_{>0}$ iff for each reactive test $T \in \mathbb{T}_{\mathrm{R,c},\phi}$ there exists a reactive test $T' \in \mathbb{T}_{\mathrm{R,c},\phi}$ such that for all sequences $\theta \in \Theta(P_1, T) \cup \Theta(P_2, T')$ of average amounts of time:

1. $prec(T, T') \geq p$ and $rec(T, T') \geq r$
2. $|prob(\mathcal{SC}^{|\theta|}_{\leq \theta \pm \epsilon, \mathcal{SC}^{|\theta|}(P_2, T')}(P_1, T)) - prob(\mathcal{SC}^{|\theta|}_{\leq \theta \pm \epsilon, \mathcal{SC}^{|\theta|}(P_1, T)}(P_2, T'))| \leq \nu$.

- Conservative extension of $\sim_{\mathrm{MT}}$.
- "Transitive".
- Checkable in poly-time.

Example

Consider P_1 and P_2 as follows:

$$<g, \gamma>.<a, \lambda + \delta>.<b, \lambda>.\underline{0} + <g, \gamma>.<a, \lambda>.<d, \lambda>.\underline{0}$$

$$<g, \gamma>.<a, \lambda>.<d', \lambda>.\underline{0} + <g, \gamma>.<a, \lambda>.<b, \lambda - \delta>.\underline{0}$$

and compare them with respect to tests whose successful computation is described by the concrete trace $g \circ a \circ *$, with $*$ any action.

Then, P_2 is Markovian testing similar to P_1 with:

- both precision and recall equal to $\frac{2}{3}$, where the difference in the observed behaviors is due to the two concrete traces $g \circ a \circ d$ of P_1 and $g \circ a \circ d'$ of P_2, under the assumption $d \neq d'$;
- temporal threshold $\epsilon \geq \frac{1}{\lambda - \delta} - \frac{1}{\lambda} > \frac{1}{\lambda} - \frac{1}{\lambda + \delta}$, where the difference in the average sojourn times is due to the three rates λ, $\lambda + \delta$, $\lambda - \delta$ labeling corresponding transitions related to the two concrete traces $g \circ a \circ b$ of P_1 and P_2;
- probability threshold 0, since the probabilities of the successful computations to compare are always the same.

Conclusions

- Testing equivalence as an ideal framework for joining two approaches (approximate behavioral equivalence vs. similarity with respect to benchmarks of typical behaviors).

- Relation with performance analysis.

- Applications to noninterference analysis.