Quantitative Aspects of Behavioural Equivalence for Real-Time Systems

Uli Fahrenberg

Department of Computer Science
Aalborg University
Denmark

May 29, 2008
1 Motivation

2 Timed traces

3 Timed languages

4 Bisimulation pseudometrics
For real-time systems and specifications, timed bisimilarity is a rather merciless concept:

The gates will be closed 1 minute before the train goes through
not timed bisimilar to
The gates will be closed 58 seconds before the train goes through

Untimed bisimilarity on the other hand is, well, useless:

The gates will be closed 1 minute before the train goes through
untimed bisimilar to
The gates will be closed 1 second before the train goes through
Motivation

- Or, using timed automata:

\[ A = \begin{array}{c}
\text{Close} \\
\rightarrow \\
x \leftarrow 0 \\
\rightarrow \\
x \geq 60 \\
\rightarrow \\
\text{Train} \\
\end{array} \]

not timed bisimilar to

\[ B = \begin{array}{c}
\text{Close} \\
\rightarrow \\
x \leftarrow 0 \\
\rightarrow \\
x \geq 58 \\
\rightarrow \\
\text{Train} \\
\end{array} \]

- And for the other case:

\[ A = \begin{array}{c}
\text{Close} \\
\rightarrow \\
x \leftarrow 0 \\
\rightarrow \\
x \geq 60 \\
\rightarrow \\
\text{Train} \\
\end{array} \]

untimed bisimilar to

\[ C = \begin{array}{c}
\text{Close} \\
\rightarrow \\
x \leftarrow 0 \\
\rightarrow \\
x \geq 1 \\
\rightarrow \\
\text{Train} \\
\end{array} \]

- Intuition: Want notion of \textit{bisimilarity up to } \varepsilon \textit{ – so that } A \sim_2 B, \text{ but } A \sim_{59} C.

- Bisimulation pseudometrics
Motivation

- Or, using timed automata:
  \[ A = \begin{array}{c}
  \xrightarrow{x \leftarrow 0} \text{Close} \\
  \xrightarrow{x \geq 60} \text{Train}
  \end{array} \]
  not timed bisimilar to
  \[ B = \begin{array}{c}
  \xrightarrow{x \leftarrow 0} \text{Close} \\
  \xrightarrow{x \geq 58} \text{Train}
  \end{array} \]

- And for the other case:
  \[ A = \begin{array}{c}
  \xrightarrow{x \leftarrow 0} \text{Close} \\
  \xrightarrow{x \geq 60} \text{Train}
  \end{array} \]
  untimed bisimilar to
  \[ C = \begin{array}{c}
  \xrightarrow{x \leftarrow 0} \text{Close} \\
  \xrightarrow{x \geq 1} \text{Train}
  \end{array} \]

- Intuition: Want notion of bisimilarity up to \( \varepsilon \) – so that \( A \sim_2 B \), but \( A \sim_{59} C \).
- Bisimulation pseudometrics
Timed traces

- Easier to define: metrics on **timed languages** (in the “linear domain”)
- Timed automata generate **timed traces**:

  \[
  L(A) = \{ (t_0, a_0, t_1, a_1, \ldots, a_n) \mid \text{exists alternating path}\]

  \[
  s_0 \xrightarrow{t_0} s'_0 \xrightarrow{a_0} s_1 \xrightarrow{t_1} s'_1 \xrightarrow{a_1} \ldots \xrightarrow{a_n} s_{n+1} \text{ in } A\}
  \]

  (In this talk, we consider only **finite** timed traces)

- **Examples**:

  \[A = \begin{align*}
  &\xrightarrow{\text{Close}} \xleftarrow{0} \xrightarrow{\geq 60} \xrightarrow{\text{Train}} \\
  \end{align*}\]

  \[L(A) = \{ (t_0, C, t_1, T) \mid t_1 \geq t_0 + 60 \} \]

  \[B = \begin{align*}
  &\xrightarrow{\text{Close}} \xleftarrow{0} \xrightarrow{\geq 58} \xrightarrow{\text{Train}} \\
  \end{align*}\]

  \[L(B) = \{ (t_0, C, t_1, T) \mid t_1 \geq t_0 + 58 \} \]

  \[C = \begin{align*}
  &\xrightarrow{\text{Close}} \xleftarrow{0} \xrightarrow{\geq 1} \xrightarrow{\text{Train}} \\
  \end{align*}\]

  \[L(C) = \{ (t_0, C, t_1, T) \mid t_1 \geq t_0 + 1 \} \]
Metrics on timed traces

- Let $\tau = (t_0, a_0, t_1, a_1, \ldots, a_n)$, $\tau' = (t'_0, a'_0, t'_1, a'_1, \ldots, a'_n)$ be two timed traces.
- If $n' \neq n$ (different length), or if $a_i \neq a'_i$ for some $i$ (difference in actions), any distance is $d(\tau, \tau') = \infty$.
- Otherwise:
  $$d_{\text{pair}}(\tau, \tau') = \max_i \{|t_i - t'_i|\}$$
  $$d_{\text{sum}}(\tau, \tau') = \max_i \{|\sum_{j=1}^i t_j - \sum_{j=1}^i t'_j|\}$$
  $$d_{\text{pair,drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{t_i}{t'_i}, \frac{t'_i}{t_i} \right) \right\} \right)$$
  $$d_{\text{sum,drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{\sum_{j=1}^i t_j}{\sum_{j=1}^i t'_j}, \frac{\sum_{j=1}^i t'_j}{\sum_{j=1}^i t_j} \right) \right\} \right)$$
Metrics on timed traces

- Let $\tau = (t_0, a_0, t_1, a_1, \ldots, a_n)$, $\tau' = (t'_0, a'_0, t'_1, a'_1, \ldots, a'_{n'})$ be two timed traces.
- If $n' \neq n$ (different length), or if $a_i \neq a'_i$ for some $i$ (difference in actions), any distance is $d(\tau, \tau') = \infty$.
- Otherwise:

$$d_{\text{pair}}(\tau, \tau') = \max_i \{|t_i - t'_i|\}$$

(measures maximal difference in pairs of delays)

$$d_{\text{sum}}(\tau, \tau') = \max_i \{|\sum_{j=1}^i t_j - \sum_{j=1}^i t'_j|\}$$

$$d_{\text{pair,drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{t_i}{t'_i}, \frac{t'_i}{t_i} \right) \right\} \right)$$

$$d_{\text{sum,drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{\sum_{j=1}^i t_j}{\sum_{j=1}^i t'_j}, \frac{\sum_{j=1}^i t'_j}{\sum_{j=1}^i t_j} \right) \right\} \right)$$
Metrics on timed traces

- Let \( \tau = (t_0, a_0, t_1, a_1, \ldots, a_n) \), \( \tau' = (t'_0, a'_0, t'_1, a'_1, \ldots, a'_{n'}) \) be two timed traces.
- If \( n' \neq n \) (different length), or if \( a_i \neq a'_i \) for some \( i \) (difference in actions), any distance is \( d(\tau, \tau') = \infty \).
- Otherwise:
  \[
d_{\text{pair}}(\tau, \tau') = \max_i \{|t_i - t'_i|\}
\]
  (measures maximal difference in pairs of delays)

  \[
d_{\text{sum}}(\tau, \tau') = \max_i \{|\sum_{j=1}^{i} t_j - \sum_{j=1}^{i} t'_j|\}
\]
  (measures maximal difference in accumulated delay)

  \[
d_{\text{pair,drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{t_i}{t'_i}, \frac{t'_i}{t_i} \right) \right\} \right)
\]

  \[
d_{\text{sum,drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{\sum_{j=1}^{i} t_j}{\sum_{j=1}^{i} t'_j}, \frac{\sum_{j=1}^{i} t'_j}{\sum_{j=1}^{i} t_j} \right) \right\} \right)
\]
Motivation
Timed traces
Timed languages
Bisimulation pseudometrics

Metrics on timed traces

Let \( \tau = (t_0, a_0, t_1, a_1, \ldots, a_n) \), \( \tau' = (t'_0, a'_0, t'_1, a'_1, \ldots, a'_{n'}) \) be two timed traces.
If \( n' \neq n \) (different length), or if \( a_i \neq a'_i \) for some \( i \) (difference in actions), any distance is \( d(\tau, \tau') = \infty \).
Otherwise:
\[
d_{\text{pair}}(\tau, \tau') = \max_i \{|t_i - t'_i|\}
\]
(measures maximal difference in pairs of delays)
\[
d_{\text{sum}}(\tau, \tau') = \max_i \{|\sum_{j=1}^i t_j - \sum_{j=1}^i t'_j|\}
\]
(measures maximal difference in accumulated delay)
\[
d_{\text{pair,drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{t_i}{t'_i}, \frac{t'_i}{t_i} \right) \right\} \right)
\]
\[
d_{\text{sum,drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{\sum_{j=1}^i t_j}{\sum_{j=1}^i t'_j}, \frac{\sum_{j=1}^i t'_j}{\sum_{j=1}^i t_j} \right) \right\} \right)
\]
(similar, but now we measure quotients (drift) instead of difference)
Metrics on timed traces

\[ d_{\text{pair}}(\tau, \tau') = \max_i \{|t_i - t'_i|\} \]
\[ d_{\text{sum}}(\tau, \tau') = \max_i \{|\sum_{j=1}^i t_j - \sum_{j=1}^i t'_j|\} \]
\[ d_{\text{pair,drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{t_j}{t'_j}, \frac{t'_j}{t_j} \right) \right\} \right) \]
\[ d_{\text{sum,drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{\sum_{j=1}^i t_j}{\sum_{j=1}^i t'_j}, \frac{\sum_{j=1}^i t'_j}{\sum_{j=1}^i t_j} \right) \right\} \right) \]

- For all of the above, \( d(\tau, \tau') = 0 \) implies \( \tau = \tau' \) (hence they are indeed metrics)
- Other metrics can be defined – e.g. with \( \sum_i \) instead of \( \max_i \)
- Most of them are topologically equivalent to one of the above (at least for finite traces)
Metrics on timed traces

\[ d_{\text{pair}}(\tau, \tau') = \max_i \{|t_i - t_i'|\} \]
\[ d_{\text{sum}}(\tau, \tau') = \max_i \{|\sum_{j=1}^i t_j - \sum_{j=1}^i t'_j|\} \]
\[ d_{\text{pair, drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{t_j}{t'_j}, \frac{t'_j}{t_j} \right) \right\} \right) \]
\[ d_{\text{sum, drift}}(\tau, \tau') = \log \left( \max_i \left\{ \max \left( \frac{\sum_{j=1}^i t_j}{\sum_{j=1}^i t'_j}, \frac{\sum_{j=1}^i t'_j}{\sum_{j=1}^i t_j} \right) \right\} \right) \]

- For all of the above, \( d(\tau, \tau') = 0 \) implies \( \tau = \tau' \) (hence they are indeed metrics)
- Other metrics can be defined – e.g. with \( \sum_i \) instead of \( \max_i \)
- Most of them are topologically equivalent to one of the above (at least for finite traces)
- (Two metrics, \( d_1 \) and \( d_2 \), are topologically equivalent iff they generate the same topology, iff there are constants \( m \) and \( M \) such that \( md_1(x, y) \leq d_2(x, y) \leq Md_1(x, y) \) for all \( x, y \)
Pseudometrics on timed languages

- For measuring differences of timed languages (which is what we want), use Hausdorff pseudometric:
  
  Given a set $X$ with pseudometric $d$, the Hausdorff pseudometric on the power set of $X$ is $d^H$ defined as follows:

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

- Hence for timed languages $L_1, L_2$ we have $d(L_1, L_2) \leq \varepsilon$ iff any timed trace in $L_1$ can be matched by a timed trace in $L_2$ with distance $\leq \varepsilon$, and vice versa – quite natural!

- So we have metrics $d_{\text{pair}}$, $d_{\text{sum}}$, $d_{\text{pair,drift}}$, $d_{\text{sum,drift}}$ for timed languages

- And $d(L_1, L_2) = 0$ iff $\text{cl } L_1 = \text{cl } L_2$, the closures of $L_1$, $L_2$ as sets of timed traces.
Pseudometrics on timed languages

Back to the examples:

\[ A = \quad \begin{array}{c}
\xrightarrow{0} \quad \xleftarrow{0} \\
\text{Close} \quad \text{Train}
\end{array} \xrightarrow{x \geq 60} C \quad \xrightarrow{t_1 \geq t_0 + 60} \]

\[ L(A) = \{ (t_0, C, t_1, T) \mid t_1 \geq t_0 + 60 \} \]

\[ B = \quad \begin{array}{c}
\xrightarrow{0} \quad \xleftarrow{0} \\
\text{Close} \quad \text{Train}
\end{array} \xrightarrow{x \geq 58} C \quad \xrightarrow{t_1 \geq t_0 + 58} \]

\[ L(B) = \{ (t_0, C, t_1, T) \mid t_1 \geq t_0 + 58 \} \]

\[ C = \quad \begin{array}{c}
\xrightarrow{0} \quad \xleftarrow{0} \\
\text{Close} \quad \text{Train}
\end{array} \xrightarrow{x \geq 1} C \quad \xrightarrow{t_1 \geq t_0 + 1} \]

\[ L(C) = \{ (t_0, C, t_1, T) \mid t_1 \geq t_0 + 1 \} \]

\[ d_{\text{pair}}(L(A), L(B)) = d_{\text{sum}}(L(A), L(B)) = 2 \]

\[ d_{\text{pair},\text{drift}}(L(A), L(B)) = d_{\text{sum},\text{drift}}(L(A), L(B)) = \log(60/58) \approx 0.015 \]

\[ d_{\text{pair}}(L(A), L(C)) = d_{\text{sum}}(L(A), L(B)) = 59 \]

\[ d_{\text{pair},\text{drift}}(L(A), L(C)) = d_{\text{sum},\text{drift}}(L(A), L(B)) = \log 60 \approx 1.8 \]
Pseudometrics on timed languages

- Back to the examples:

\[ A = \quad \text{Close} \xrightarrow{\text{Train}} \xrightarrow{\geq 60} \quad \text{Train} \]

\[ L(A) = \{(t_0, C, t_1, T) \mid t_1 \geq t_0 + 60\} \]

\[ B = \quad \text{Close} \xrightarrow{\text{Train}} \xrightarrow{\geq 58} \quad \text{Train} \]

\[ L(B) = \{(t_0, C, t_1, T) \mid t_1 \geq t_0 + 58\} \]

\[ C = \quad \text{Close} \xrightarrow{\text{Train}} \xrightarrow{\geq 1} \quad \text{Train} \]

\[ L(C) = \{(t_0, C, t_1, T) \mid t_1 \geq t_0 + 1\} \]

\[ d_{\text{pair}}(L(A), L(B)) = d_{\text{sum}}(L(A), L(B)) = 2 \]

\[ d_{\text{pair,drift}}(L(A), L(B)) = d_{\text{sum,drift}}(L(A), L(B)) = \log(60/58) \approx 0.015 \]

\[ d_{\text{pair}}(L(A), L(C)) = d_{\text{sum}}(L(A), L(B)) = 59 \]

\[ d_{\text{pair,drift}}(L(A), L(C)) = d_{\text{sum,drift}}(L(A), L(B)) = \log 60 \approx 1.8 \]

- Problem: for timed automata \( A, B \), it is **undecidable** whether \( L(A) = L(A) \), hence all our pseudometrics on timed languages are most probably **uncomputable** in general!
Bisimulation pseudometrics

- Back to the “branching domain”: It is **decidable** whether two timed automata are **bisimilar**

  \[ \Rightarrow \text{Want to introduce bisimulation pseudometrics on timed automata which correspond to these pseudometrics on timed languages} \]

- **correspond** should mean: \[ d(A, B) = \varepsilon < \infty \iff d(L(A), L(B)) = \varepsilon \]

- **in other words**: For automata with finite bisimulation distance, the language mapping should be **distance-preserving**.
Bisimulation pseudometrics

- **Pair version:** For states $s_1, s_2$ in timed transition systems $A, B$, say that $s_1 \sim_{\varepsilon}^{\text{pair}} s_2$ iff

\[
\forall s_1 \xrightarrow{a} s'_1 \in T_1 : \exists s_2 \xrightarrow{a} s'_2 \in T_2 : s'_1 \sim_{\varepsilon}^{\text{pair}} s'_2 \\
\land \forall s_2 \xrightarrow{a} s'_2 \in T_2 : \exists s_1 \xrightarrow{a} s'_1 \in T_1 : s'_1 \sim_{\varepsilon}^{\text{pair}} s'_2 \\
\land \forall s_1 \xrightarrow{t_1} s'_1 \in T_1 : \exists s_2 \xrightarrow{t_2} s'_2 \in T_2 : s'_1 \sim_{\varepsilon}^{\text{pair}} s'_2 \land |t_1 - t_2| \leq \varepsilon \\
\land \forall s_2 \xrightarrow{t_2} s'_2 \in T_2 : \exists s_1 \xrightarrow{t_1} s'_1 \in T_1 : s'_1 \sim_{\varepsilon}^{\text{pair}} s'_2 \land |t_1 - t_2| \leq \varepsilon
\]

(Recall that for timed traces, $d_{\text{pair}}(\tau, \tau') = \max_i \{|t_i - t'_i|\}$ )

- Define $d_{\text{pair}}(A, B) = \inf\{\varepsilon \mid A \sim_{\varepsilon}^{\text{pair}} B\}$

- Then the $L$ mapping is indeed distance-preserving

- Similar can be done for $d_{\text{pair,drift}}$

- What about computability?
Bisimulation pseudometrics

- The sum version is more difficult: Need to remember differences in delays across transitions
- For states $s_1$, $s_2$ in timed transition systems $A$, $B$, say that $s_1 \sim_{\text{sum}}^\varepsilon \delta s_2$ iff

\[
\forall s_1 \xrightarrow{a} s'_1 \in T_1 : \exists s_2 \xrightarrow{a} s'_2 \in T_2 : s'_1 \sim_{\varepsilon, \delta} \text{sum} \ s'_2 \\
\wedge \forall s_2 \xrightarrow{a} s'_2 \in T_2 : \exists s_1 \xrightarrow{a} s'_1 \in T_1 : s'_1 \sim_{\varepsilon, \delta} \text{sum} \ s'_2 \\
\wedge \forall s_1 \xrightarrow{t_1} s'_1 \in T_1 : \exists s_2 \xrightarrow{t_2} s'_2 \in T_2 : s'_1 \sim_{\varepsilon, \delta + t_1 - t_2} \text{sum} \ s'_2 \wedge |\delta + t_1 - t_2| \leq \varepsilon \\
\wedge \forall s_2 \xrightarrow{t_2} s'_2 \in T_2 : \exists s_1 \xrightarrow{t_1} s'_1 \in T_1 : s'_1 \sim_{\varepsilon, \delta + t_1 - t_2} \text{sum} \ s'_2 \wedge |\delta + t_1 - t_2| \leq \varepsilon
\]

($\delta$ is the lead which $A$ hitherto has worked up compared to $B$)
- Define $d_{\text{sum}}(A, B) = \inf\{\varepsilon \mid A \sim_{\varepsilon}^\text{sum} B\}$ as before
- This is work by Henzinger, Majumdar, Prabhu (FORMATS 2005)
- (Similar can be done for $d_{\text{sum,drift}}$)
- Yes, the $L$ mapping is again distance-preserving
- And HMP’05 shows that $d_{\text{sum}}$ is computable!
Workshop on Approximate Behavioural Equivalences

ABE 08, the Workshop on Approximate Behavioural Equivalences, will take place at the University of Toronto on Monday August 18, 2008. The workshop is affiliated with CONCUR 08.