Quantitative Aspects of Behavioural Equivalence for Real-Time Systems

Uli Fahrenberg

Department of Computer Science
Aalborg University
Denmark

May 29, 2008

Motivation

- 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:

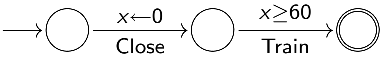
  The gates will be closed 1 minute before the train goes through

  untimed bisimilar to

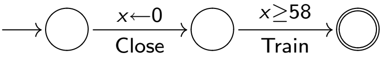
  The gates will be closed 1 second before the train goes through

Motivation

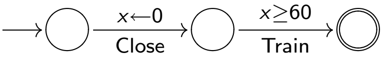
- Or, using timed automata:

$A =$ 

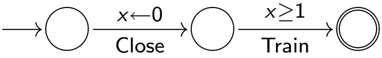
not timed bisimilar to

$B =$ 

- And for the other case:

$A =$ 

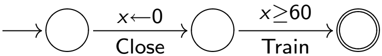
untimed bisimilar to

$C =$ 

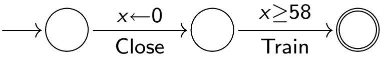
- Intuition: Want notion of bisimilarity up to ε – so that $A \sim_2 B$, but $A \sim_{59} C$.
- Bisimulation metrics

Motivation

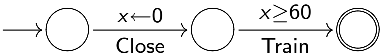
- Or, using timed automata:

$A =$ 

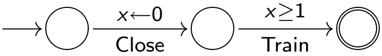
not timed bisimilar to

$B =$ 

- And for the other case:

$A =$ 

untimed bisimilar to

$C =$ 

- Intuition: Want notion of bisimilarity up to ε – so that $A \sim_2 B$, but $A \sim_{59} C$.
- Bisimulation pseudometrics

Timed traces

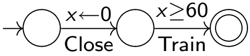
- Easier to define: metrics on timed languages (in the "linear domain")
- Timed automata generate timed traces:

$$L(A) = \big\{ (t_0, a_0, t_1, a_1, \dots, a_n) \;\big|\; \text{exists alternating path } s_0 \xrightarrow{t_0} s_0' \xrightarrow{a_0} s_1 \xrightarrow{t_1} s_1' \xrightarrow{a_1} \dots \xrightarrow{a_n} s_{n+1} \text{ in } A \big\}$$

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

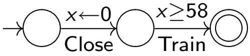
- Examples:

$A =$



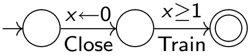
$$L(A) = \big\{ (t_0, \mathrm{C}, t_1, \mathrm{T}) \;\big|\; t_1 \ge t_0 + 60 \big\}$$

$B =$



$$L(B) = \big\{ (t_0, \mathrm{C}, t_1, \mathrm{T}) \;\big|\; t_1 \ge t_0 + 58 \big\}$$

$C =$



$$L(C) = \big\{ (t_0, \mathrm{C}, t_1, \mathrm{T}) \;\big|\; t_1 \ge t_0 + 1 \big\}$$

Metrics on timed traces

- Let $\tau = (t_0, a_0, t_1, a_1, \dots, a_n)$, $\tau' = (t'_0, a'_0, t'_1, a'_1, \dots, a'_{n'})$ 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:

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

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

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

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

Metrics on timed traces

- Let $\tau = (t_0, a_0, t_1, a_1, \dots, a_n)$, $\tau' = (t'_0, a'_0, t'_1, a'_1, \dots, a'_{n'})$ 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:

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

(measures maximal difference in pairs of delays)

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

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

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

Metrics on timed traces

- Let $\tau = (t_0, a_0, t_1, a_1, \dots, a_n)$, $\tau' = (t'_0, a'_0, t'_1, a'_1, \dots, a'_{n'})$ 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:

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

(measures maximal difference in pairs of delays)

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

(measures maximal difference in accumulated delay)

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

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

Metrics on timed traces

- Let $\tau = (t_0, a_0, t_1, a_1, \dots, a_n)$, $\tau' = (t'_0, a'_0, t'_1, a'_1, \dots, a'_{n'})$ 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:
$$d_{\text{pair}}(\tau, \tau') = \max_i \left\{ |t_i - t'_i| \right\}$$

(measures maximal difference in pairs of delays)

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

(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

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

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

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

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

- 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

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

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

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

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

- 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 $md_1(x,y) \le d_2(x,y) \le Md_1(x,y)$ for all x, y)

Pseudometrics on timed languages

- For measuring differences of timed languages (which is what we want), use Hausdorff pseudometric:

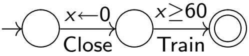
  Given a set X with pseudometric d, the Hausdorff pseudometric on the power set of X is d^{H} defined as follows:

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

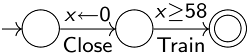
- Hence for timed languages L_1, L_2 we have $d(L_1, L_2) \le \varepsilon$ iff any timed trace in L_1 can be matched by a timed trace in L_2 with distance $\le \varepsilon$, and vice versa – quite natural!
- So we have metrics d_{pair}, d_{sum}, $d_{\text{pair,drift}}$, $d_{\text{sum,drift}}$ for timed languages
- And $d(L_1, L_2) = 0$ iff $\operatorname{cl} L_1 = \operatorname{cl} L_2$, the closures of L_1, L_2 as sets of timed traces.

Pseudometrics on timed languages

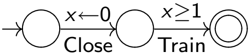
- Back to the examples:

$A =$ 

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

$B =$ 

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

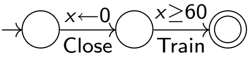
$C =$ 

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

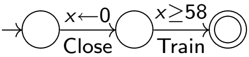
$$d_{\text{pair}}\big(L(A), L(B)\big) = d_{\text{sum}}\big(L(A), L(B)\big) = 2$$
$$d_{\text{pair,drift}}\big(L(A), L(B)\big) = d_{\text{sum,drift}}\big(L(A), L(B)\big) = \log(60/58) \approx .015$$
$$d_{\text{pair}}\big(L(A), L(C)\big) = d_{\text{sum}}\big(L(A), L(B)\big) = 59$$
$$d_{\text{pair,drift}}\big(L(A), L(C)\big) = d_{\text{sum,drift}}\big(L(A), L(B)\big) = \log 60 \approx 1.8$$

Pseudometrics on timed languages

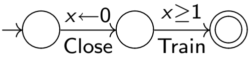
- Back to the examples:

$A =$ 

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

$B =$ 

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

$C =$ 

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

$$d_{\text{pair}}\big(L(A), L(B)\big) = d_{\text{sum}}\big(L(A), L(B)\big) = 2$$
$$d_{\text{pair,drift}}\big(L(A), L(B)\big) = d_{\text{sum,drift}}\big(L(A), L(B)\big) = \log(60/58) \approx .015$$
$$d_{\text{pair}}\big(L(A), L(C)\big) = d_{\text{sum}}\big(L(A), L(B)\big) = 59$$
$$d_{\text{pair,drift}}\big(L(A), L(C)\big) = d_{\text{sum,drift}}\big(L(A), L(B)\big) = \log 60 \approx 1.8$$

- 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

$\Longrightarrow$ Want to introduce bisimulation pseudometrics on timed automata which correspond to these pseudometrics on timed languages

- correspond should mean: $d(A, B) = \varepsilon < \infty \Longrightarrow d\big(L(A), L(B)\big) = \varepsilon$
- 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

$$
\begin{aligned}
&\forall s_1 \xrightarrow{a} s_1' \in T_1 : \exists s_2 \xrightarrow{a} s_2' \in T_2 : s_1' \sim_\varepsilon^{\text{pair}} s_2' \\
\wedge\ &\forall s_2 \xrightarrow{a} s_2' \in T_2 : \exists s_1 \xrightarrow{a} s_1' \in T_1 : s_1' \sim_\varepsilon^{\text{pair}} s_2' \\
\wedge\ &\forall s_1 \xrightarrow{t_1} s_1' \in T_1 : \exists s_2 \xrightarrow{t_2} s_2' \in T_2 : s_1' \sim_\varepsilon^{\text{pair}} s_2' \wedge |t_1 - t_2| \le \varepsilon \\
\wedge\ &\forall s_2 \xrightarrow{t_2} s_2' \in T_2 : \exists s_1 \xrightarrow{t_1} s_1' \in T_1 : s_1' \sim_\varepsilon^{\text{pair}} s_2' \wedge |t_1 - t_2| \le \varepsilon
\end{aligned}
$$

- (Recall that for timed traces, $d_{\text{pair}}(\tau, \tau') = \max_i \big\{ |t_i - t_i'| \big\}$)
- 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

$$\begin{aligned}
&\forall s_1 \xrightarrow{a} s_1' \in T_1 : \exists s_2 \xrightarrow{a} s_2' \in T_2 : s_1' \sim^{\text{sum}}_{\varepsilon,\delta} s_2' \\
\wedge\, &\forall s_2 \xrightarrow{a} s_2' \in T_2 : \exists s_1 \xrightarrow{a} s_1' \in T_1 : s_1' \sim^{\text{sum}}_{\varepsilon,\delta} s_2' \\
\wedge\, &\forall s_1 \xrightarrow{t_1} s_1' \in T_1 : \exists s_2 \xrightarrow{t_2} s_2' \in T_2 : s_1' \sim^{\text{sum}}_{\varepsilon,\delta+t_1-t_2} s_2' \wedge |\delta + t_1 - t_2| \le \varepsilon \\
\wedge\, &\forall s_2 \xrightarrow{t_2} s_2' \in T_2 : \exists s_1 \xrightarrow{t_1} s_1' \in T_1 : s_1' \sim^{\text{sum}}_{\varepsilon,\delta+t_1-t_2} s_2' \wedge |\delta + t_1 - t_2| \le \varepsilon
\end{aligned}$$

- (δ is the lead which A hitherto has worked up compared to B)
- Define $d_{\text{sum}}(A, B) = \inf\{\varepsilon \mid A \sim^{\text{sum}}_{\varepsilon} B\}$ 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_{sum} is computable!

Advertisement

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.