Models are cool!

In Computer Science, Formal models are useful for

- understanding computational aspects
- design
- analysis
- verification

Hundreds of computational concepts $\Rightarrow$ hundreds of different models

- (dozens of) automata, Petri nets, process algebras, ...

Each comes with a rich theory of results, proof techniques, tools...

Metamodels are cooler!

I like *meta*models!

A good metamodel is useful insomuch as it provides

- unifying mathematical theory of many models
- general results, logics and tools, which can be readily instantiated
- cross-fertilizing connections between models
- scenario for comparing models
- deeper insights

Metamodels for Quantitative Systems

Several metamodels have been proposed (**WLTS**, **ULTraS**, **FuTS**)

- covers many kinds of quantitative models (non-determistic probabilistic, stochastic, timed . . .).
- provides a general definition of bisimilarity
- general results about strong quantitative bisimulation [M. & Peressotti, QAPL'14]

But other observational equivalences for quantitative systems (weak, trace, branching, delay. . .) are not as well understood as strong bisimulation.

- unobservable actions may have observable effects (e.g., execution times, probabilities, energy consumption)
- not a single definition, but many "ad hoc"
- sometimes, no agreement on what is the "right" definition
- no clear categorical characterization

. . . the perfect situation where a metamodel can be useful.

Weak bisimulations for labelled transition systems with quantitative aspects

Marino Miculan[1] Marco Peressotti

Dept. of Mathematics, Computer Science and Physics, University of Udine
Dept. of Mathematics and Computer Science, University of Southern Denmark

iFM[2], 2019.04.05

---

[1] marino.miculan@uniud.it

[2] it is just an exponent.

In this talk: *weak weighted bisimulation*

We give a general definition of *weak bisimulation* valid for a wide range of labelled transition systems, namely *LTS weighted over semirings*.

1. general: it encompasses many known systems
2. decidable: a uniform algorithm applicable to various semirings
3. with a categorical coalgebraic construction.

Applications:

- obtaining weak bisimulations and decision algorithms for new kinds of systems
- generalize further to other classes of systems (beyond weighted LTS) and to other behavioural equivalences (beyond weak bisimilarity)

Weighted Labelled Transition Systems

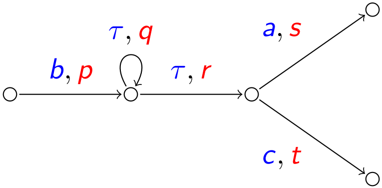
Let $\mathfrak{W} = (W, +, 0)$ be a commutative monoid.

Definition ([Klin, 2009])

A *($\mathfrak{W}$-weighted) labelled transition system* is a triple (X, A, ρ) where:

- X is a set of *states* (processes);
- A is a set of *labels* (actions);
- $\rho : X \times A \times X \to W$ is a *weight function*.

Transitions can be thought to be labelled with actions and weights drawn from $\mathfrak{W}$, with the unit 0 disabling transitions.



Weighted Labelled Transition Systems

Let $\mathfrak{W} = (W, +, 0)$ be a commutative monoid.

Definition ([Klin, 2009])

A *($\mathfrak{W}$-weighted) labelled transition system* is a triple (X, A, ρ) where:

- X is a set of *states* (processes);
- A is a set of *labels* (actions);
- $\rho : X \times A \times X \to W$ is a *weight function*.

Different $\mathfrak{W}$ yield different systems and bisimulation:

- usual non-deterministic LTS: $2 = (\{\mathtt{tt}, \mathtt{ff}\}, \vee, \mathtt{ff})$;
- stochastic LTS: $(\mathbb{R}_0^+, +, 0)$
- fully probabilistic LTS: $(\mathbb{R}_0^+, +, 0)$ such that $\forall x : \sum_{a,y} \rho(x \xrightarrow{a} y) \in \{0, 1\}$
- *etc.*

Weighted (strong) bisimulation

Definition ([Klin, 2009])

A *(strong)* $\mathfrak{W}$*-bisimulation* on (X, A, ρ) is an equivalence relation $R \subseteq X \times X$ such that $(x, x') \in R$ iff for each label $a \in A$ and each equivalence class C of R:

$$\sum_{y \in C} \rho(x \xrightarrow{a} y) = \sum_{y \in C} \rho(x' \xrightarrow{a} y).$$

Using different $\mathfrak{W}$ we can recover different systems and bisimulation:

- $(\{\mathtt{tt}, \mathtt{ff}\}, \vee, \mathtt{ff})$: strong non-deterministic bisimulation (Milner);
- $(\mathbb{R}_0^+, +, 0)$: strong stochastic bisimulation (Hillstone, Panangaden);
- $(\mathbb{R}_0^+, +, 0)$: strong probabilistic bisimulation (Larsen-Skou);
- *etc.*

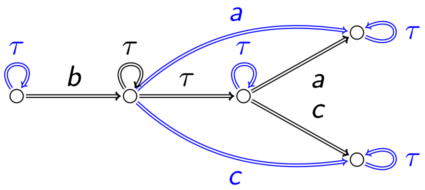
Weak bisimulation: the non-deterministic case via "double arrow" construction

Definition ([Milner, ages ago])

$R \subseteq X \times X$ is a *weak bisimulation on* $(X, A + \{\tau\}, \longrightarrow)$ iff for each $(x, x') \in R$, label $\alpha \in A + \{\tau\}$ and equivalence class $C \in X/R$:

$$\exists y \in C. x \stackrel{\alpha}{\Longrightarrow} y \iff \exists y' \in C. x' \stackrel{\alpha}{\Longrightarrow} y'$$

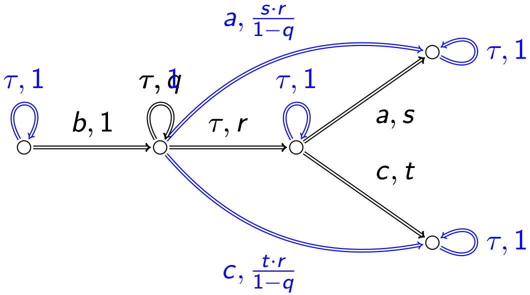
where $\Longrightarrow \subseteq X \times (A \uplus \{\tau\}) \times X$ is the τ-reflexive-transitive closure of $\longrightarrow$.



$\approx$ for $(X, A + \{\tau\}, \longrightarrow)$ is $\sim$ for $(X, A + \{\tau\}, \Longrightarrow)$.

Generalizing the non-deterministic case?

What if we apply the same approach to a fully-probabilistic system ($\sum \rho \in 0, 1$)?



This is *not* a probabilistic system, and the bisimulation induced is *not* a weak probabilistic bisimulation (in the sense of Baier-Hermanns).

We need a better definition of *saturation*

Weak bisimulation: the fully-probabilistic case

Definition ([Baier-Hermanns, 97])

$R \subseteq X \times X$ is a *weak (probabilistic) bisimulation* on $(X, A + \{\tau\}, \mathrm{P})$ iff for $(x, x') \in R$, $a \in A$ and equivalence class $C \in X/R$:

$$\mathrm{Prob}(x, \tau^* a \tau^*, C) = \mathrm{Prob}(x', \tau^* a \tau^*, C)$$
$$\mathrm{Prob}(x, \tau^*, C) = \mathrm{Prob}(x', \tau^*, C).$$

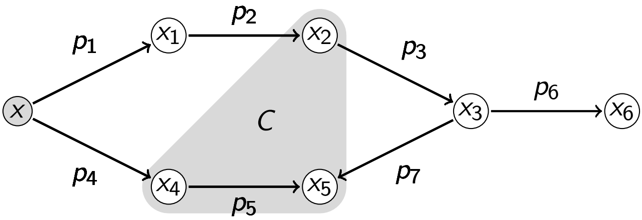
where Prob is the extension over finite execution paths of the unique probability measure induced by P.

Intuitively...

$\mathrm{Prob}(x, T, C)$ is the probability of **reaching** C from x generating some trace in T.

States of C cannot be considered separately because σ-additivity does not hold (i.e. $\mathrm{Prob}(x, T, C_1 \cup C_2) \neq \mathrm{Prob}(x, T, C_1) + \mathrm{Prob}(x, T, C_2)$)

τ-closure vs. reachability: probabilistic

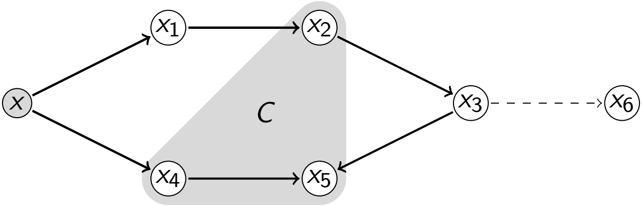


Assuming p_i is the probability of an action, what is the probability to reach class C from x?

$$\mathbf{1} \between (p_1 \cdot p_2) + (p_4) + (p_4 \cdot p_5) + (p_1 \cdot p_2 \cdot p_3 \cdot p_7)$$

(we ignored labels, but can be easily taken into account).

τ-closure vs. reachability: non-deterministic



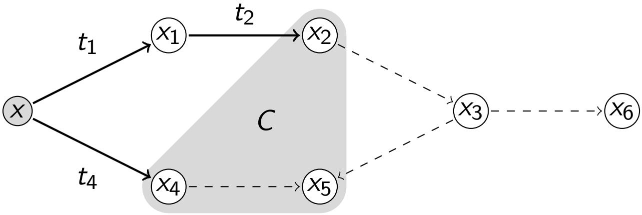
Assuming the non-deterministic case ($p_i = \mathtt{tt}$), can we reach C from x?

$$\mathtt{tt} = (\mathtt{tt} \wedge \mathtt{tt}) \vee (\mathtt{tt}) \vee (\mathtt{tt} \wedge \mathtt{tt}) \vee (\mathtt{tt} \wedge \mathtt{tt} \wedge \mathtt{tt} \wedge \mathtt{tt})$$

Here τ-closure and reachability coincide...

But this is a very specific case (and for a *precise* reason.)

τ-closure vs. reachability: stochastic



Assuming t_i describes the time consumed by an action, how much time takes to go from x to C?

$$t = \min(t_1 + t_2, t_4)$$

Weighting execution paths

Previous examples used two operations on weights:

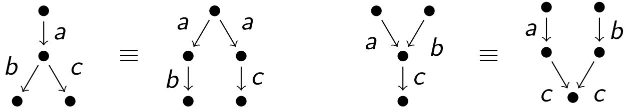
- $(W, +, 0)$ for **branching** (a commutative monoid)
- $(W, \cdot, 1)$ for **chaining** (a monoid)

Subject to some coherence conditions:

- 0 expresses **termination** (annihilates chaining)

$$0 \cdot a = 0 = a \cdot 0$$

- **independence** of execution paths



$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \qquad (a + b) \cdot c = (a \cdot c) + (b \cdot c)$$

Semirings of weights

Henceforth, let $\mathfrak{W} = (W, +, 0, \cdot, 1)$ be a **semiring** (*cf.* $\mathfrak{W}$-automata).

Definition (Path weight)

Given a weight function ρ, its extension to finite paths is:

$$\rho(x_0 \xrightarrow{a_1} x_1 \dots \xrightarrow{a_n} x_n) \triangleq \rho(x_0 \xrightarrow{a_1} x_1) \cdot \ldots \cdot \rho(x_{n-1} \xrightarrow{a_n} x_n)$$

Weighting finite paths is enough for our aims since two (countably) infinite paths are observationally distinguished iff there is a finite path telling them apart *i.e.* by finite observation.

(Countably infinite paths require countable multiplication, or equivalently a sufficiently expressive notion of limits).

Weighting sets of paths

It is enough to weight (particular) sets of paths.

Definition (Finite paths to C)

For a state x, a set of traces T and a set of states C, the set of finite paths reaching C from x with trace in T is

$$\wr x, T, C \wr \triangleq \left\{ \pi \in \mathrm{FPaths}(x) \;\middle|\; \begin{array}{l} \mathrm{last}(\pi) \in C,\ \mathrm{trace}(\pi) \in T, \\ \forall \pi' \preceq \pi : \mathrm{trace}(\pi') \in T \Rightarrow \mathrm{last}(\pi') \notin C \end{array} \right\}$$

(This corresponds to a particular "saturation" construction in the coalgebraic setting)

Weak $\mathfrak{W}$-bisimulation

Definition (Weak $\mathfrak{W}$-bisimulation)

$R \subseteq X \times X$ is a *weak $\mathfrak{W}$-bisimulation* for $(X, A + \{\tau\}, \rho)$ iff for all $(x, x') \in R$, $a \in A$ and equivalence class $C \in X/R$, the following hold:

$$\rho(\wr x, \tau^*, C \int) = \rho(\wr x', \tau^*, C \int)$$
$$\rho(\wr x, \tau^* a \tau^*, C \int) = \rho(\wr x', \tau^* a \tau^*, C \int).$$

Remark

- Weak $\mathfrak{W}$-bisimulation is just categorical weak bisimulation, concretely presented in the case of WLTS.
- Other bisimulations can be obtained by changing the set of paths (e.g., for delay bisimulation: $\wr x, \tau^*, C \int$ and $\rho(\wr x, \tau^* a, C \int)$)

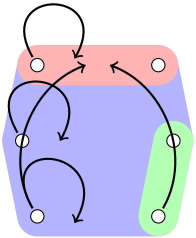
Examples of weak $\mathfrak{W}$-bisimulation

- Non-deterministic systems and Milner's weak bisimulation: *Boolean semiring*: $(\{\mathtt{tt}, \mathtt{ff}\}, \vee, \mathtt{ff}, \wedge, \mathtt{tt})$
- Fully-probabilistic systems and Baier-Hermanns's weak bisimulation:
  - *Positive real semiring*: $(\overline{\mathbb{R}}_0^+, +, 0, \cdot, 1)$
  - *Probabilistic σ-semiring*: $([0,1], +, 0, \cdot, 1)$
- Stochastic systems (and a new weak bisimulation): *transition-time random variables semiring*: $\mathfrak{S} \triangleq (\mathbb{T}, \min, \mathcal{T}_{+\infty}, +, \mathcal{T}_0)$
- Troubleshooting: *Likelihood semiring*: $([0,1], \max, 0, \cdot, 1)$
- Optimization problems (especially scheduling):
  - *Tropical semiring*: $(\overline{\mathbb{R}}_0^+, \min, +\infty, +, 0)$
  - *Arctic semiring*: $(\overline{\mathbb{R}}, \max, -\infty, +, 0)$
  - *Bottleneck semiring*: $(\overline{\mathbb{R}}_0^+, \min, +\infty, \max, 0)$
- Formal languages: *Free language semiring*: $(\wp(\Sigma^*), \cup, \emptyset, \circ, \varepsilon)$
- And many more...

Deciding Weak Weighted Bisimulation

Computing weak $\mathfrak{W}$-bisimulation

We generalize Kanellakis-Smolka's algorithm for strong bisimulation of *finite* LTSs [Kanellakis-Smolka 1989].
Let $(X, A + \{\tau\}, \rho)$ be a finite $\mathfrak{W}$-LTS and let P be a partition of X.



$$P_0 = \{X\}$$

$$\rho(\wr x_0, \tau^* a \tau^*, X \wr) = \rho(\wr x_1, \tau^* a \tau^*, X \wr)$$

General algorithm for weak weighted bisimulation

$\mathcal{X} \leftarrow \{X\}$
$\mathcal{X}' \leftarrow \emptyset$
repeat
    changed $\leftarrow$ **false**
    $\mathcal{X}'' \leftarrow \mathcal{X}$
    for all $C \in \mathcal{X} \setminus \mathcal{X}'$ **do**
        for all $\sigma \in \Sigma + \{\tau\}$ **do**
            if $\langle \sigma, C \rangle$ is a splitter **then**
                $\mathcal{X} \leftarrow \bigcup \{B/\underset{\sigma,C}{\approx} \mid B \in \mathcal{X}\}$
                changed $\leftarrow$ **true**
            end if
        end for
    end for
    $\mathcal{X}' \leftarrow \mathcal{X}''$
until not *changed*
return $\mathcal{X}$

Computing the weight of redundancy-free sets

Question

Given x, a, C, how do we compute $\rho(\wr x, \tau^*, C \int)$ and $\rho(\wr x, \tau^* a \tau^*, C \int)$?

By solving a system of linear equations over $\mathfrak{W}$.
For each state x, let x_τ, x_a be two variable over $\mathfrak{W}$.
Equations:

$$x_\tau = \begin{cases} 1 & \text{if } x \in C \\ \sum_{y \in X} \rho(x, \tau, y) \cdot y_\tau & \text{otherwise} \end{cases}$$

$$x_a = \sum_{y \in X} \rho(x, a, y) \cdot y_\tau + \sum_{y \in X} \rho(x, \tau, y) \cdot y_a$$

Intuition: $x_\tau = \rho(\wr x, \tau^*, C \int)$ $\qquad x_a = \rho(\wr x, \tau^* a \tau^*, C \int)$

Solvability of the equation systems

The definitions of x_a's form a linear equation system $x = A \cdot x + b$, which defines an operator over W^n (A is $n \times n$).

$$F(y) = A \cdot y + b$$

Proposition

If S is ω-complete and ω-continuous then the system has a unique solution.

A semiring is called

- positive (or zero-sum-free) if $v + w = 0 \implies v = w = 0$;
- ω-complete if it is positive and has countable sums $\sum_{i<\omega} w_i = \sup\{\sum_{j \in J} w_j \mid J \subset \omega\}$;
- ω-continuous if suprema of ω-chains exist and are preserved by both operations.

Examples: all semirings in this talk.

Specific domain solvers can be used as "plugins" of the algorithm.

Complexity

Almost same complexity of Kanellakis-Smolka's original algorithm, but:

- No constant-time random-access data structures;
- No pre-computed (saturated) transitions.

Proposition (Time complexity)

The asymptotic upper bound for time complexity of the proposed algorithm is in

$$\mathcal{O}(nm(\mathcal{L}_W(n) + n^2))$$

where $n = |X|$ *and* $m = |A| + 1$ *and* $\mathcal{L}_W(n)$ *is the time complexity of solving a system of* n *linear equations with* n *variables over semiring* W*.*

If binary joins of W-matrices left-distributes over matrix multiplication, saturation can be precomputed. Time complexity: $\mathcal{O}(\mathcal{L}_W(n) + mn^2)$.

In presence of constant-time random-access data structures time complexity is in $\mathcal{O}(nm(\mathcal{L}_W(n) + n))$

Conclusions

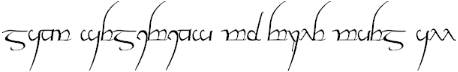
Done:

- framework for defining strong and weak bisimilarities (and beyond);
- semirings where saturation is effectively computable;
- general algorithm, parametric in the semiring (and solver); may be not as efficient as ad hoc algorithms, but still is a first implementation
- categorical coalgebraic characterization (not shown).

To do:

- real implementation
- generalizing Paige-Tarjan instead of Kanellakis-Smolka
- symbolic weighted bisimulation

Thanks for your attention.



(Many semirings to rule them all.)