ENS DE LYON

Codensity bisimulations for Weighted Finite Automata

Work supervised by Hasuo Ichiro in collaboration with Eberhart Clovis and Komorida Yuichi

HUMEAU Samuel

August 26, 2019

Abstract

We use the codensity bisimilarity framework from [8] to retreive two notions of bisimulation on weighted finite automata. Doing so we give a coalgebraic representation of weighted finite automata, and a formulation of the two bisimulation notions by greatest-fixed point of well-chosen functions. We show those notions are equivalent to the original ones in [1] and to the ones we get from the framework.

I Introduction

The idea of bisimulation is really important when considering a system. It has been first introduced by Park and Milner in [9, 10]. For example, two states of a Kripke frame are bisimilar if and only if they satisfy the same formulas of the modal logic (see the Hennessy-Milner theorem in [2]). Intuitively, two states are bisimilar, if their behavior, i.e. what can be seen from outside, are exactly the same. In this sense it defines a semantics. Often it is represented by an equivalence relation. The greatest bisimulation is called a bisimilarity relation.

Most of the time, there is an associated game to a bisimilarity relation (many examples can be found in the main reference of this work [8]). The game is played between Spoiler and Duplicator. In it, Duplicator asserts that two states are bisimilar, Spoiler that they are not. These games should be defined in such a way that Duplicator wins if and only if the states the game starts from are bisimilar. The advantage of introducing games is that they sometimes come with algorithms to decide bisimilarity.

Recently there has been a rise of interest for quantitative bisimulation (starting with [3]). In this context, a bisimulation is not a relation anymore but a pseudo-metric on the state space. The pseudo-metric bisimilarity is the least pseudo-metric bisimulation, point-wise. Its kernel must coincide with the usual notion of relation bisimilarity when it exists. Defining bisimulation as a metric stresses the fact that some states, while not bisimilar, are very close behavior-wise. In this work we will be interested in both bisimilarity relations and quantitative bisimilarity. We will often say metric instead of pseudo-metric.

Let us look at the simple example of Kripke frames. A Kripke frame is just an unlabelled transition system, i.e., states with a relation:

Definition

A *Kripke frame* is a pair (X, R) where X is a set, called *state space* and $R \subseteq X \times X$ is a relation on X called *transition relation*. If $(x, x') \in R$, then it is said that there is a *one step computation* from x to x'. We will note xRx' in this case.

Note that a Kripke frame on the state space X can equivalently be characterized as a transtition function $c\colon X \to \mathcal{P}(X)$. This is a coalgebraic representation which has proved to be well-suited (see [7]). Finding the right characterization is a common starting point of categorical studies of systems.

For many types of system the bisimilarity relation, or the behavior, is defined locally, through observables (the state being blocking or not for Kripke frame, accepting or not for deterministic finite automata), and the behavior of the successor states. Two states are bisimilar if their observables are the same, and their successors are again bisimilar. It looks like an induction when considering two states. Still it is not when constructing the whole relation. The bisimilarity relation of many systems can be expressed as the greatest fixed point of a well-chosen function. This function express the local understanding we have of what a bisimulation is. Later we will present the codensity bisimilarity framework (see [8]). This framework is based on a greatest fixed point definition of bisimilarity. We will try to focus on it.

In a Kripke frame, two states are bisimilar if and only if they have the same observables and their successor states are again bisimilar. This defines a function which takes a relation and change it to a new relation containing those states that have the same observables and whose successor states are related by the original relation. A fixed point of this function will be a bisimulation. Take (X, R) a Kripke frame and $\psi_R\colon \mathcal{P}(X \times X) \to \mathcal{P}(X \times X)$ defined by:

$$\forall B \in \mathcal{P}(X \times X),\ \psi_R(B) = \left\{ (x, y) \in X^2 \,\middle|\, \begin{cases} \forall x' \in X,\ xRx' \Rightarrow (\exists y' \in X,\ yRy' \wedge x'By') \\ \forall y' \in X,\ yRy' \Rightarrow (\exists x' \in X,\ xRx' \wedge x'By') \end{cases} \right\}$$

Then:

Proposition 1. *ψ_R has a greatest fixed point.*

Proof. This will be done using the Knaster-Tarski theorem: if $f\colon L \to L$ acts on a complete lattice L and is order preserving, then the set of fixed points of f also has a lattice structure. Here, $(\mathcal{P}(X \times X), \subseteq)$ is a complete lattice, and ψ_R is obviousy order preserving. ψ_R has a greatest fixed point.

The bisimilarity relation of Kripke frame (X, R) is defined as the greatest fixed point of ψ_R.

Categorical framework have been widely used in recent years to understand systems, logic, and unify similar notions in different contexts ([7, 4, 6, 5]). This work consists in the application of one of those framework, the codensity bisimilarity one ([8], based on [11]) to a particular example, the one of weighted finite automata, noted WFA. The way bisimulation will be treated for WFA before applying the framework is very similar to what is done with Kripke frame here.

Objectives

The goal of this work is to apply the codensity bisimilarity framework from [8] to weighted finite automata (a certain type of automaton) with linear and quantitative bisimulations ([1]). The framework defines an abstract notion of bisimilation, of bisimilarity, and of associated games. It ensures good properties between the games and the bisimulation notion. It uses a coalgebraic representation of systems, and it defines bisimilarity as the greatest fiexd-point of a particular functor.

Weighted finite automata are DFA, except that state space is a vector space, and transition functions are linear functions. The definition of the two bisimulations notion used in this work are find in [1], but they are ill-suited for the codensity bisimilarity framework. The first step of our work will be to have a coalgebraic definition of WFA, and then to characterise the two bisimulations in terms of fixed points in order to make them fit the framework.

Their is no proof this framework can be applied to recover known notions of bisimilarity for all types of system. Understanding how powerful it is already is of interest. One of our goals was thus to test how expressive this framework is by trying to recover more complex notions of bisimulations than the ones it had already been applied to. Furthermore, as this framework allows an automatic definition of games associated to the bisimilarity, it could lead to new algorithms to compute bisimilarity on WFAs, or approximation algorithms.

Moreover the codensity bisimilarity framework allows for a lot of leeway in how it can be applied (for example in the choice of its so-called *modalities*), wih may lead to new notions of bisimulations. This is future work.

Outline

The section II will introduce different notions needed to reach the objectives. Category theory tools used for the codensity bisimilarity framework will be defined, the framework's direct background presented, and two defintions of WFA given and the main part of the framework explained in two versions. The section [**?**] and [**?**] will concentrate on retreiving two bisimulation notion on WFAs using the said framework. The final section [**?**] will conclude.

II Backgrounds

This section will introduce the different tools needed to present the results on WFA: category theory definitions, codensity bisimilarity framework, and weighted finite automata.

1/ Category theory

We assume the reader is familiar with the definitions of category, functor, and is a bit used to categorical reasoning. Here we need two definitions, coalgebra and fibrations. Coalgebras are one of the main categorical frameworks for the study of systems. It is in particular well-suited for the study of bisimulations [7]. It represents the system itself. Fibrations can be seen as a way to organize some predicates on the state space.

Definition

Let $\mathbb{C}$ a category and $F\colon \mathbb{C} \to \mathbb{C}$ an endofunctor. An *F-coalgebra* or just coalgebra, is given by an object $X \in \mathbb{C}$ along with a morphism $c\colon X \to F(X)$. X is usually called the state space, and c the transition function.

To a state $x \in X$, c associates $c(x) \in F(X)$. The idea is that X appears in $F(X)$ as the succesor of x, the other elements appearing in $F(X)$ are side information about the system.

Recall that in our example a Kripke frame was represented by a function $c\colon X \to \mathcal{P}(X)$. This is a $\mathcal{P}$-coalgebra on the category *Set* of sets and functions. Here there is no side information. That is because the observables is to know if the state is blocking or not, i.e. if $c(x) = \emptyset$ or not. Deterministic finite automata are represented by coalgebras of the type $c\colon X \to X^{\Sigma} \times 2$, where $2 = \{0, 1\}$ and Σ is the alphabet. The first part is the transition function. It requires a letter to make the transition. Here there is side information. 2 will evaluate to 1 if the state is accepting, and to 0 if it is not.

Fibrations require the definition of cartesian morphism:

Definition

Let $\mathbb{E}$ and $\mathbb{B}$ two categories, $p\colon \mathbb{E} \to \mathbb{B}$ a functor, and $\overline{f}\colon e_U \to e_V$ in $\mathbb{E}$. We will note $f\colon U \to V$ for $p(\overline{f}) = f$. $\overline{f}$ is a cartesian morphism if for all $g : W \to U$ in $\mathbb{B}$, and $\overline{h} : e_W \to e_V$ such that $p(\overline{h}) = f \circ g$ and $p(e_W) = W$, there is a unique morphism $\overline{g} : e_W \to e_U$ and $\overline{h} = \overline{f} \circ \overline{g}$.

$$
\begin{tikzcd}
 & e_W \arrow[rr, "\overline{f \circ g}"] \arrow[dr, dashed, "\overline{g}"'] & & e_V \\
\mathbb{E} \arrow[dd, "p"'] & & e_U \arrow[ur, "\overline{f}"'] & \\
 & & & \\
\mathbb{B} & W \arrow[rr, "f \circ g"] \arrow[dr, "g"'] & & V \\
 & & U \arrow[ur, "f"'] &
\end{tikzcd}
$$

Figure 1: The diagram of the cartesian property

Fibration are used to structure bisimulation candidates. It is defined on a category $\mathbb{E}$, generally represented above the category $\mathbb{C}$. If $X \in \mathbb{C}$, there is a category above X, called fiber, and noted $\mathbb{E}_X$. It is a subcategory of $\mathbb{E}$. In our case, it is a complete lattice above X. If $P \in \mathbb{E}_X$, it asserts "P is a bisimulation" and thus can be seen as a predicate.

Definition

Let $\mathbb{B}$ and $\mathbb{E}$ two categories and $p\colon \mathbb{E} \to \mathbb{B}$ a functor. $\mathbb{B}$ is called the base category, and $\mathbb{E}$ the total category. If $X \in \mathbb{B}$ and $e_X \in \mathbb{E}$ such that $p(e_X) = X$, e_X is said to be *above* X. If $f\colon U \to V$ in $\mathbb{B}$ and $\overline{f}\colon e_U \to e_V$ in $\mathbb{E}$ such that $p(\overline{f}) = f$, then $\overline{f}$ is said to be above f.
We say that $p\colon \mathbb{E} \to \mathbb{B}$ is a fibration if, given $f\colon U \to V$ in $\mathbb{B}$ and e_V above V in $\mathbb{E}$, there is an object $f^* e_V$ above U and a morphism $\overline{f}\colon f^* U \to e_V$ above f that is a cartesian morphism.
Take $X \in \mathbb{B}$. The collection of all objects above X along with all the morphisms above the identity Id_X is called the *fiber* over X and is noted $\mathbb{E}_X$. It is a subcategory of $\mathbb{E}_X$.

Proposition 2. *Let $p\colon \mathbb{E} \to \mathbb{B}$ a fibration and $f\colon U \to V$ a morphism in $\mathbb{B}$.*

Define $f^*\colon \mathbb{E}_V \to \mathbb{E}_U$ as the operation that sends $e_V \in \mathbb{E}_V$ to $f^*e_V \in \mathbb{E}_U$ as defined above. We call this operation pullback. It can be extended to morphisms, making it a functor.

If $m\colon e_V^1 \to e_V^2$ in $\mathbb{E}_V$, f^*m is given by universal property of cartesian morphism as in:

$$
\begin{tikzcd}
 & f^*Y_1 \arrow[rr, "m \circ \overline{f_1}"] \arrow[dr, dashed, "f^*m"'] & & Y_2 \\
\mathbb{E} \arrow[d, "p"'] & & f^*Y_2 \arrow[ur, "\overline{f_2}"'] & \\
\mathbb{B} & U \arrow[rr, "f \circ Id_U"] \arrow[dr, "Id_U"'] & & V \\
 & & U \arrow[ur, "f"'] &
\end{tikzcd}
$$

Figure 2: Property of cartesian morphism applied to f

In our context some particular properties are required:

Definition

A $\text{Clat}_\sqcap$-fibration is a fibration $p\colon \mathbb{E} \to \mathbb{B}$ such that for all $X \in \mathbb{B}$, $\mathbb{E}_X$ is a complete lattice, and such that all pullback functors f^* are $\text{Clat}_\sqcap$-morphisms, i.e. monotone morphisms that preserve all meets. The associated order is called *indistinguishability order* and will be noted $\sqsubseteq$. An element $P \in \mathbb{E}$ is called indistinguishability predicate.

The idea is that the indistinguishability order is not necessarily the usual order: $P \sqsubseteq Q$ is supposed to express the fact that in Q states look more bisimilar than in P. For relation bisimulation, and object of $\mathbb{E}$ is a relation, or an equivalence relation. The order is the usual one, the inclusion order. $P \sqsubseteq Q$ means that everything considered bisimilar by P is also considered bisimilar by Q, but Q can contain more bisimilar states than P. For metric bisimulation, the object of $\mathbb{E}$ are pseudo-metrics. The usual order is the point-wise one. Here the order will be inversed. $P \sqsubseteq Q$ means that states are closer, or more bisimilar, using Q than they are using P.

Definition

Let $p\colon \mathbb{E} \to \mathbb{B}$ a $\text{Clat}_\sqcap$-fibration. A morphism $f\colon U \to V$ in $\mathbb{B}$ is said *decent* from P above U to Q above V if there is a unique arrow $\overline{f}\colon P \to Q$ above f.

2/ Codensity bisimulation

Three parameters are essential in a bisimulation. One is an object that is used to compare the observables, and is noted Ω. It can be thought as a set of truth values. It is given with the second one, a way of comparing the different truth values that is compatible with the bisimulation notion. It is noted $\boldsymbol{\Omega}$ and will be in the fiber of Ω. For example the equality relation if we want the observables to be the same, and if we are talking about relation bisimulation. The last one is called modality, it is a morphism of type $\tau\colon F\Omega \to \Omega$, with F the coalgebra type. It is used to accomodate truth values with the transition structure. Those parameters are needed to understand the bisimulation notion locally.

Let us look at DFA. They are represented by a coalgebra $c\colon X \to X^\Sigma \times 2$ where X is the state space, Σ the alphabet and $2 = \{0,1\}$ expresses the fact the state is accepting or not. The observables is being accepting or not. This will be compared in $\Omega = 2$. For relational bisimulation, having the same observables is reasonable. 2 is given with $\mathbf{\Omega} = Eq_2$ to compare the observables. After a transition when working with truth values, we have something of the type $F(2) = 2^\Sigma \times 2$. We would like a function that takes this and give back a value in 2. This is the modality. For DFA it is the must modality: every value must be 1, meaning that for each letter, the successor states of the two states under comparison must be bisimilar again.

The three parameters are the starting point of the bisimilarity codensity framework that will be presented here. Then, the framework defines a codensity lifting of the functor F. F is defined in the base category of a fibration and it is lifted to the total category.

Definition

Take a functor $F\colon \mathbb{C} \to \mathbb{C}$ and a $\mathbf{Clat}_\sqcap$-fibration $p\colon \mathbb{E} \to \mathbb{C}$. The *codensity lifting* $F^{\mathbf{\Omega},\tau}\colon \mathbb{E} \to \mathbb{E}$ along p with parameters $(\mathbf{\Omega}, \tau)$ where

- $\Omega \in \mathbb{C}$ is the set of truth values.
- $\tau\colon F\Omega \to \Omega$ is the modality (this is an F-algebra).
- $\mathbf{\Omega}$ above Ω is the observation domain.

is the endofunctor on $\mathbb{E}$ defined by

$$F^{\mathbf{\Omega},\tau} P = \bigsqcap_{k \in \mathbb{E}(P,\mathbf{\Omega})} (\tau \circ F(p(k)))^* \,\mathbf{\Omega}$$

on objects. On morphisms it is defined by mapping $l\colon P \to Q$ to the unique arrow above $F(p(l))$ that goes from $F^{\mathbf{\Omega},\tau}P$ to $F^{\mathbf{\Omega},\tau}Q$ on arrows.

The definition of the functor on morphisms uses the fact that $F(p(l))$ is decent (see [8]).

Often multiple parameters are needed. Each set of parameters is used for one computation type. For DFA, there is actually one modality for each letter, and one for the observables. The codensity lifting of functor becomes:

Definition

Let a family of parameters $\{\mathbf{\Omega}_a, \tau_a\}_{a \in \mathbb{A}}$ indexed by a class $\mathbb{A}$. The codensity lifting is defined as:

$$F^{\mathbf{\Omega},\tau} P = \bigsqcap_{a \in \mathbb{A}} F^{\mathbf{\Omega}_a,\tau_a} P = \bigsqcap_{a \in \mathbb{A}} \bigsqcap_{k \in \mathbb{E}(P,\mathbf{\Omega})} (\tau \circ F(p(k)))^* \,\mathbf{\Omega}$$

Defining bisimulation through a lifting of functors is a well-known methods (see [7] chapter 3). Here the lifting allows both the definition of a predicate transformer whose greatest-fixed point will be the bisimilarity, and of associated games.

Definition

$P \in \mathbb{E}_X$ is an $(\mathbf{\Omega}, \tau)$-bisimulation over the F-coalgebra $c\colon X \to FX$ if c is decent from P to $F^{\mathbf{\Omega},\tau}P$.
A predicate transformer $\Phi_c^{\mathbf{\Omega},\tau}\colon \mathbb{E}_X \to \mathbb{E}_X$ is defined by:

$$\Phi_c^{\mathbf{\Omega},\tau} P = c^*(F^{\mathbf{\Omega},\tau} P)$$

Finaly the codensity bisimilarity is the greatest fixed point of the predicate transformer and is noted $\nu\Phi_c^{\mathbf{\Omega},\tau}$.

Some good properties in [8] justify those definitions. For example P is a codensity bisimulation over c if and only if $P \sqsubseteq \Phi_c^{\mathbf{\Omega},\tau} P$. For relation this expresses the fact that a bisimulation is just an invariant predicate of the right predicate transformers. This is a very much wanted property.

Kripke frames only requires one parameter. As said above, $\Omega = \{0,1\}$. The modality used is the must one; $\tau\colon \mathcal{P}\{0,1\} \to \{0,1\}$ is defined as:

$$\begin{aligned}\tau(\emptyset) &= 1\\ \tau(\{0\}) &= 0\\ \tau(\{1\}) &= 1\\ \tau(\{0,1\}) &= 0\end{aligned}$$

The $\mathrm{Clat}_\sqcap$-fibration goes from *EqRel*, the category of equivalence relations on sets, with relation preserving functions as morphisms, to *Set* the category of sets and functions. On objects it associates its underlying set to a relation. On morphisms it is the identity. The fiber on X is the category of equivalence relation on X. The order in the fiber is the inclusion one. The bisimilarity relation this fibration defines through the framework is the same as the one presented in introduction.

A great part of [8] is not presented here. The main interest of this framework is that once the right parameters is defined, both the bisimulation and an associated equivalent game can be derived from it, with good properties ensured. Finding parameters is thought to be simpler than coming up with the right game. Applying this part of the framework to our examples with WFA is the next thing after this work.

3/ Weighted Finite Automata

Weighted finite automata are very similar to deterministic finite automata: each transition requires a letter, but transition follow linear functions, a state being a point in a vector space. There are more general definitions of WFA, defined with semirings, but here we will restrict to WFAs on real finite vector spaces.

Definition

A WFA is a quintuple $\left(V, \Sigma, (\tau_\sigma)_{\sigma\in\Sigma}, \alpha, \beta\right)$, where V is a vector space, Σ is an alphabet, τ_σ for $\sigma \in \Sigma$ are linear applications, $\alpha \in V$ is the initial state and $\beta\colon V \to \mathbb{R}$ is a linear form corresponding to the final weight.

This is the definition given in [1]. The first step of our work is to give a coalgebraic characterization of WFA.

Proposition 3. *A weighted finite automata can be given by a coalgebra* $c\colon V \overset{\tau_{(-)},\beta}{\to} V^\Sigma \times \mathbb{R}$ *on the category of vectorial spaces, where* Σ *is an alphabet and* $(\tau_\sigma)_{\sigma\in\Sigma}$ *is a* Σ*-indexed linear transition functions family.*

Through the notations the correspondance is straightforward. The only missing thing in the second definition is the initial state. It is implicit. We did not add it as it is non relevant to the rest of this work.

Definition

The *computation* $\tau_\sigma\colon V \to V$ of c on $x \in V$ along $w_1 \cdots w_n = w \in \Sigma^*$ is given inductively by:

$$\begin{aligned}\tau_\varepsilon(x) &= x \text{ with } \varepsilon \text{ the empty word}\\ \tau_w(x) &= \tau_{w_n}(\tau_{w_1\ldots w_{n-1}}(x))\end{aligned}$$

A *computation path* is an element of $\Sigma^\infty = \Sigma^* \cup \Sigma^\mathbb{N}$ the set of finite and infinite words on Σ.

Now we are going to look at bisimulation notions for WFAs. For each of those we will give the definition from [1], give our characterization as greatest fixed point, and show how we applied the codensity bisimilarity framework.

III Linear bisimulation

Definition and characterization

Definition

Given $c\colon V \overset{\tau_{(-)},\beta}{\to} V^\Sigma \times \mathbb{R}$ a WFA, a bisimulation subspace is a subspace $W \subseteq V$ such that:

- For all $x \in W$, $\beta(x) = 0$.
- W is invariant under every τ: for all $\sigma \in \Sigma$, $\tau_\sigma(W) \subseteq W$.

$(x, y) \in V^2$ are said bisimilar if their exists a bisimulation subspace W such that $x - y \in W$.

This notion of bisimulation distinguishes states using the obervable function β. Two states must be bisimilar if and only if the same real number is observed on them with β, and for each computation type, i.e. for each letter, the associated new states are again bisimilar. This gives a local undertanding of what a bisimulation is. Using this, define $\psi\colon \mathcal{P}(V \times V) \to \mathcal{P}(V \times V)$ by:

$$(x, y) \in \psi(S) \Leftrightarrow \beta(x - y) = 0 \wedge \forall \sigma \in \Sigma,\ (\tau_\sigma(x), \tau_\sigma(x)) \in S$$

ψ has a greatest fixed point. This is still done using the Knaster-Tarski theorem with $\mathcal{P}(V \times V)$ as complete lattice and ψ as order-preserving function. It is pretty straightforward.

Definition

The bisimilarity relation $\sim_c$ on c is defined to be the greatest fixed point of ψ.

It is natural to look for the following property:

Lemma 1. *$x \sim_c y$ if and only if their is a bisimulation subspace containing $x - y$.*

In order to prove this, we will use the following characterization:

Proposition 4. *The greatest fixed point of ψ is equal to:*

$$\Delta = \{(x, y) | \forall w \in \Sigma^*,\ \beta(\tau_w(x)) = \beta(\tau_w(y))\}$$

meaning that the same observations are made along any computation path started on x and y.

Proof. First, let us show Δ is a fixed point of ψ:

$$\begin{aligned}
(x, y) \in \psi(\Delta) \Leftrightarrow & \beta(x - y) = 0 \wedge \forall \sigma \in \Sigma,\ (\tau_\sigma(x), \tau_\sigma(y)) \in \Delta \\
\Leftrightarrow & \beta(x) = \beta(y) \wedge (\forall \sigma \in \Sigma,\ \forall w \in \Sigma^*,\ \beta(\tau_w(\tau_\sigma(x))) = \beta(\tau_w(\tau_\sigma(y)))) \\
\Leftrightarrow & \beta(x) = \beta(y) \wedge (\forall \sigma \in \Sigma,\ \forall w \in \Sigma^*,\ \beta(\tau_{\sigma.w}(x)) = \beta(\tau_{\sigma.w}(y))) \\
\Leftrightarrow & \beta(x) = \beta(y) \wedge \forall w \in \Sigma^{\geq 1},\ \beta(\tau_w(x)) = \beta(\tau_w(y)) \\
\Leftrightarrow & \forall w \in \Sigma^*,\ \beta(\tau_w(x)) = \beta(\tau_w(y)) \\
\Leftrightarrow & (x, y) \in \Delta
\end{aligned}$$

Δ is a fixed point of ψ. Thus, it is included in the greatest fixed point of ψ.

Suppose (x, y) is in the greatest fixed point of ψ. Let $w \in \Sigma^*$. Let us prove by induction on w's length that $(\tau_w(x), \tau_w(y)) \in \nu\psi$. If $|w| = 0$, $\tau_w(x) = x$, $\tau_w(y) = y$, and $(\tau_w(x), \tau_w(y)) \in \nu\psi$. Suppose $|w| = n + 1$ for some $n \in \mathbb{N}$, and that for any $\sigma \in \Sigma^*$ such that $|\sigma| \leq n$, $(\tau_w(x), \tau_w(y)) \in \nu\psi$. Then, w can be written $w = w'.\alpha$ with $\alpha \in \Sigma$ and $|w'| \leq n$. By induction hypothesis $(\tau_{w'}(x), \tau_{w'}(y)) \in \nu\psi = \psi(\nu\psi)$. Thus, by definition of ψ, for any $\sigma \in \Sigma$, $(\tau_\sigma(\tau_{w'}(x)), \tau_\sigma(\tau_{w'}(y))) \in \nu\psi$. This can be rewritten, with $\sigma = \alpha$ as $(\tau_{w'.\alpha}(x), \tau_{w'.\alpha}(y)) \in \nu\psi$, i.e. $(\tau_w(x), \tau_w(y)) \in \nu\psi$.

By induction, if $(x, y) \in \nu\psi$ then for any $w \in \Sigma^*$ we have $(\tau_w(x), \tau_w(y)) \in \nu\psi$.

Now this means that if $(x, y) \in \nu\psi$, then for any $w \in \Sigma^*$ $(\tau_w(x), \tau_w(y)) \in \psi(\nu\psi)$, i.e. that for any $w \in \Sigma^*$, $\beta(\tau_w(x)) = \beta(\tau_w(y))$, and that $(x, y) \in \Delta$. So,

$$\nu\psi \subseteq \Delta$$

and finally,

$$\nu\psi = \Delta$$

□

Now let us prove the proposition:

Proof. Suppose $x - y$ is in some bisimulation subspace W. As W is invariant by τ_σ for any $\sigma \in \Sigma$, then it also is invariant under τ_w for any $w \in \Sigma^*$, meaning that $\beta(x) = \beta(y)$ as $x - y \in W$ but also that for any $w \in \Sigma^*$, $\beta(\tau_w(x)) = \beta(\tau_w(y))$ as $\tau_w(x - y) \in W$ by invariance. Thus, (x, y) is in the greatest fixed point of ψ using the lemma 1.
Now, let us show that the set made by the differences $x - y$ of pairs $(x, y) \in \nu\psi$ of this greatest fixed point is a bisimulation subspace. This set will be noted B. Applying ψ one time gives that for any $(x, y) \in \nu\psi$, $\beta(x) = \beta(y)$. If $\delta \in B$ there exists $(x, y) \in \nu\psi$ such that $\delta = x - y$. Then $\beta(\delta) = 0$ as $\beta(\delta) = \beta(x - y)$ and $\beta(x) = \beta(y)$. $\nu\psi$ being a fixed point of ψ, by definition, for any $\sigma \in \Sigma$ and $(x, y) \in \nu\psi$, $(\tau_\sigma(x), \tau_\sigma(y)) \in \nu\psi$, i.e. for any $\delta \in B$, $\tau_\sigma(\delta) \in B$. Thus, $\tau_\sigma(B) \subseteq B$.

Thus, if B is a linear subspace of V, then it is a bisimulation subspace.

Let $\delta_1, \delta_2 \in B$ and $\lambda \in \mathbb{R}$. There exists, by definition of B, $x, y, z, w \in V$ such that $(x, y) \in \nu\psi$, $(w, z) \in \nu\psi$ and $\delta_1 = x - y$, $\delta_2 = w - z$. Then $\delta_1 + \lambda\delta_2 = (x + \lambda w) - (y + \lambda z)$. Take a word $\sigma \in \Sigma^*$. Then

$$\begin{aligned}\beta(\tau_\sigma(x + \lambda w)) &= \beta(\tau_\sigma(x)) + \lambda\beta(\tau_\sigma(w)) \\ &= \beta(\tau_\sigma(y)) + \lambda\beta(\tau_\sigma(z)) \\ &= \beta(\tau_\sigma(y + \lambda z))\end{aligned}$$

By the above characterization of $\nu\psi$, $(x + \lambda w, y + \lambda z) \in \nu\psi$, implying that $\delta_1 + \lambda\delta_2 \in B$. B is a bisimulation subspace, ending the proof. □

Codensity bisimilarity framework Here we will apply the codensity framework to recover this notion of linear bisimulation. The goal is to compute the above codensity lifting from which a notion of codensity bisimilarity and codensity bisimilarity game can be derived using the right parameters. The bisimilarity notion should coincide with the one presented above.

As a reference, the detailed steps are the following:

- Define the category we are in and recall the form of the functor F we are interested in, i.e. representing the right kind of systems.
- Define the right $\mathbf{Clat}_\sqcap$-fibration. Its fibers should represent indistinguishability predicates on the associated objects.
- Define Ω the object representing the truth values.
- Define $\tau \colon F(\Omega) \to \Omega$ a modality representing how to "flatten" the truth values generated by F's type.
- Define $\boldsymbol{\Omega}$ an object above Ω representing the kind of predicates on the truth values we are interested in to define our bisimulation notion.
- Work out the codensity lifting.
- Work out the codensity bisimilarity. Show it defines the right notion.

A bisimulation should define an equivalence relation. But here it is not enough. The first definition of linear bisimulation we gave used bisimulation subspace. Relations defining a bisimulation should follow the structure existing on the state space. Congruences on V, i.e. equivalence relations stable by linear operations, will be our indistinguishability predicates. The categories will be the following.

Definition

Define $Cong(V)$ the category of congruences on the real finite vector space V with inclusion order as morphisms.
Define $Cong$ the category of congruences on real finite vector spaces with linear functions $f\colon V \to W$ such that if $(x, y) \in R$ then $(f(x), f(y)) \in S$ as morphism from $R \in Cong(V)$ to $S \in Cong(W)$.
$Cong(V)$ is a subcategory of $Cong$.

If $f\colon V \to W$ is the linear function underlying a morphism from R to S in $Cong$, it will be noted $\overline{f}$ when considered in $Cong$.

We need to show they are categories:

Proof. $Cong(V)$ is a particular case of poset. As such it is a category. $Cong$ contains objects and morphisms. Identities obviously exists, and composition respects the constraints on morphisms: if $f\colon V \to W$ and $g\colon W \to U$ are two linear applications underlying morphisms from R to S and from S to T, then $g \circ f\colon V \to U$ is again a linear application and if $(x, y) \in R$, then $(f(x), f(y)) \in S$, and as such, $(g(f(x)), g(f(y))) \in T$. Composition is obviously associative as it is induced by the usual one on linear functions. Thus, $Cong$ contains objects and morphisms, its morphisms have an associative composition structure, each object has an identity. So $Cong$ is a category. □

Systems are modeled by coalgebras of the form $c\colon V \to V^{\Sigma} \times \mathbb{R}$ on the category of real finite vector spaces. The corresponding functor will be noted F. To an object V it associates $V^{\Sigma} \times \mathbb{R}$. To a morphism $f\colon V \to W$ it associates the morphisms from $F(V)$ to $F(W)$ given by $f(\tau_{(-)}) \times Id_{\mathbb{R}}$.

Before looking at the codensity lifting we still need to define the $\mathrm{Clat}_{\sqcap}$-fibration and the right parameters.

Proposition 5. *The functor $p\colon Cong \to Vect_{\mathbb{R}}$ which associates V to a congruence relation in $Cong(V)$ and f to an arrow $\overline{f}\colon R \to S$ in $Cong$ where $R \in Cong(V)$ $S \in Cong(W)$, and f is the linear morphism from V to W underlying the arrow in $Cong(\mathbb{R})$, is a $Clat_{\sqcap}$ fibration.*

Proof. p is a functor. It is well-defined on objects and on morphisms. If $\overline{f}\colon R \to S$ with $R \in Cong(V)$ and $S \in Cong(W)$, then there is $f\colon V \to W$ underlying $\overline{f}$ and $p(\overline{f}) = f$, meaning $p(\overline{f})\colon p(R) \to p(S)$. p has the right form. As the composition in $Cong$ is defined using the composition on $Vect_{\mathbb{R}}$, $p(\overline{f} \circ \overline{g}) = p(\overline{f \circ g}) = f \circ g$. By definition of p, $p(Id_R) = Id_V = Id_{p(R)}$. Thus p is indeed a functor.

Let us show it is a fibration. Let V and W two real linear spaces, $f\colon V \to W$ a linear function, and S in $Cong$ above W. Define $f^*S = \{(x, y) | (f(x), f(y)) \in S\}$. It is an object of $Cong(V)$. Then by definition, f appears as a morphism from f^*S to W. This gives the cartesian lifting of f. If the lifting has the following universal property, then p is a fibration: for all $g\colon U \to V$ in $Vect(\mathbb{R})$, and $\overline{h}\colon R \to S$ in $CongEq$ above $f \circ g$, there is a unique $\overline{g}\colon R \to f^*S$ above g such that $\overline{h} = \overline{f} \circ \overline{g}$. The morphism $\overline{g}$ is uniquely determined, by definition of $Cong$, to be g. What needs to be shown is that g suits the situation. Suppose $(x, y) \in R$. Then $(h(x), h(y)) \in S$, meaning $(f(g(x)), f(g(y))) \in S$. By definition of f^*S, this means $(g(x), g(y)) \in f^*S$. Thus, g is indeed a morphism from R to f^*S in $Cong$. By definition of composition in $Cong$, $\overline{h} = \overline{f} \circ \overline{g}$. Thus, p is a fibration.
Let us show it is a $\mathrm{Clat}_{\sqcap}$-fibration. Each fiber $Cong(V)$ is defined to be a complete lattice, with $\emptyset$ and $V \times V$ as bottom and top elements, $\cap$ as meet, and $\oplus$ as join. Remains to show the pullback functors preserve all meets. Here meets are simply intersections. Take f a linear function from V to W, $(S_i)_{i \in I}$ above W. Then,

$$\begin{aligned} f^*\left(\bigcap_{i \in I} S_i\right) &= \left\{(x, y) \in V^2 | \forall i \in I,\ (f(x), f(y)) \in S_i\right\} \\ &= \bigcap_{i \in I} \left\{(x, y) \in V^2 | (f(x), f(y)) \in S_i\right\} \\ &= \bigcap_{i \in I} f^*(S_i) \end{aligned}$$

Thus pullbacks preserves all meets. p is a $\mathrm{Clat}_{\sqcap}$-fibration. □

The systems we are interested in are WFA. Similar to DFA they read words during their computations. Following what we said in 2/ we will use multiple parameters. We associates one modality to each letter, plus one to the observables. For every $a \in \Sigma$ we define $\Omega_a = \mathbb{R}$ and $\tau_a \colon \Omega_a^\Sigma \times \mathbb{R} \to \Omega_a$ the projection function using $a \in \Sigma$, and $\Omega_\beta = \mathbb{R}$ with $\tau_\beta \colon F(\Omega_\beta) \to \Omega_\beta$ the projection on the observables.

States should be compared in $\mathbb{R}$. That is why all Ω are set to $\mathbb{R}$. Projections are just a way of distinguishing the different computation types. Before computing the codensity lifting and codensity bisimilarity, one thing is missing: $\mathbf{\Omega}$. $\mathbf{\Omega}$ should represent how two values in Ω define bisimilarity. As $\Omega = \mathbb{R}$ has been chosen because observation are done in $\mathbb{R}$, and as two states should be bisimilar if and only if the same observations can be done on both of them, and on any computation path, we chose $\mathbf{\Omega} = Eq_\Omega$. Now let us compute the codensity lifting.

Proposition 6. *With above notation and $P \in Cong(V)$:*

$$F^{\Omega,\tau}P = \left\{((\sigma_1, o_1), (\sigma_2, o_2)) \in \left(V^\Sigma \times \mathbb{R}\right)^2 | o_1 = o_2 \wedge (\forall a \in \Sigma,\ \sigma_1(a) P \sigma_2(a))\right\}$$

Proof. In a multiple observation domains context, the codensity lifting $F^{\Omega,\tau} \colon Cong \to Cong$ is defined by, $P \in Cong(V)$:

$$\begin{aligned}
F^{\Omega,\tau}P &= \bigsqcap_{a \in \Sigma \cup \{\beta\}} \bigsqcap_{k \in Cong(P, \mathbf{\Omega}_a)} (\tau_a \circ F(p(k)))^* \,\mathbf{\Omega}_a \\
&= \bigcap_{a \in \Sigma \cap \{\beta\}} \bigcap_{\substack{k \colon V \to \mathbb{R} \text{ linear} \\ (x,y) \in \mathbb{P} \Rightarrow k(x) = k(y)}} (\tau_a \circ F(k))^* \, Eq_\mathbb{R} \\
&= \bigcap_{a \in \Sigma \cap \{\beta\}} \bigcap_{\substack{k \colon V \to \mathbb{R} \text{ linear} \\ (x,y) \in \mathbb{P} \Rightarrow k(x) = k(y)}} \left\{(x, y) \in F(V)^2 | (\tau_a \circ F(k))(x) = (\tau_a \circ F(k))(y)\right\} \\
&= \bigcap_{a \in \Sigma \cap \{\beta\}} \bigcap_{\substack{k \colon V \to \mathbb{R} \text{ linear} \\ (x,y) \in \mathbb{P} \Rightarrow k(x) = k(y)}} \left\{((o, \sigma), (o', \sigma')) \in (V^\Sigma \times \mathbb{R})^2 | (\tau_a \circ F(k))(o, \sigma) = (\tau_a \circ F(k))(o', \sigma')\right\} \\
&= (\bigcap_{a \in \Sigma} \bigcap_{\substack{k \colon V \to \mathbb{R} \text{ linear} \\ (x,y) \in \mathbb{P} \Rightarrow k(x) = k(y)}} \left\{((o, \sigma), (o', \sigma')) \in (V^\Sigma \times \mathbb{R})^2 | k(\sigma(a)) = k(\sigma'(a))\right\}) \\
&\qquad \cap (\left\{((o, \sigma), (o', \sigma')) \in \left(V^\Sigma \times \mathbb{R}\right)^2 | o = o'\right\}) \\
&= \bigcap_{\substack{k \colon V \to \mathbb{R} \text{ linear} \\ (x,y) \in \mathbb{P} \Rightarrow k(x) = k(y)}} (\bigcap_{a \in \Sigma} \left\{((o, \sigma), (o', \sigma')) \in (V^\Sigma \times \mathbb{R})^2 | k(\sigma(a)) = k(\sigma'(a))\right\}) \\
&\qquad \cap (\left\{((o, \sigma), (o', \sigma')) \in \left(V^\Sigma \times \mathbb{R}\right)^2 | o = o'\right\}) \\
&= \bigcap_{\substack{k \colon V \to \mathbb{R} \text{ linear} \\ (x,y) \in \mathbb{P} \Rightarrow k(x) = k(y)}} \left\{((o, \sigma), (o', \sigma')) \in (V^\Sigma \times \mathbb{R})^2 | \forall a \in \Sigma k(\sigma(a)) = k(\sigma'(a))\right\} \\
&\qquad \cap (\left\{((o, \sigma), (o', \sigma')) \in \left(V^\Sigma \times \mathbb{R}\right)^2 | o = o'\right\}) \\
&= \bigcap_{\substack{k \colon V \to \mathbb{R} \text{ linear} \\ (x,y) \in \mathbb{P} \Rightarrow k(x) = k(y)}} \left\{((o, \sigma), (o', \sigma')) \in (V^\Sigma \times \mathbb{R})^2 | o = o' \wedge \forall a \in \Sigma k(\sigma(a)) = k(\sigma'(a))\right\}
\end{aligned}$$

Take $(x, y) \in V^2$ such that there exists $(x_1, y_1), \cdots, (x_n, y_n) \in (V^2)^n$, $x - y = x_1 + \cdots + x_n - y_1 - \cdots - y_n$, and for all $i \in \{1, \cdots, n\}$, $x_i P y_i$. Then, $x = x_1 + \cdots + x_n - y_1 - \cdots - y_n + y$ so that by reflexitivity, $xP(x_1 + \cdots + x_n - y_1 - \cdots - y_n + y)$. But P being a congruence, $x_1 + \cdots + x_n P y_1 + \cdots + y_n$, and by reflexitivity and substraction, $0_V P x_1 + \cdots + x_n - y_1 - \cdots - y_n$ and finally, xPy.

Now, define $k \colon V \to \mathbb{R}$ linear such that its kernel is exactly the subspace generated by $x - y$ when xPy. If $x - y$ is in its kernel, i.e. $k(x) = k(y)$, necessarily, xPy using what we just showed. Using this k for $a \in \Sigma$, $k(\sigma(a)) = k(\sigma'(a))$ if and only if $\sigma(a)P\sigma'(a)$, showing that: Then:

$$F^{\Omega,\tau}P = \left\{((o,\sigma),(o,\sigma')) \in (V^{\Sigma} \times \mathbb{R})^2 | o = o' \wedge \forall a \in \Sigma,\ \sigma(a)P\sigma'(a)\right\}$$ □

Now we would like to check that this codensity lifting gives rise to the same notion of bisimilarity than the bisimilarity relation $\sim_c$, i.e. the bisimulation subspace reformulated as arising from fixed points.

Proposition 7. *The predicate transformer $\Phi_c^{\mathbf{\Omega},\tau}$ is given by, $c = (c_\beta, (c_a)_{a\in\Sigma})$:*

$$\begin{aligned}\Phi_c^{\mathbf{\Omega},\tau}P &= c^*(F^{\mathbf{\Omega},\tau}P)\\ &= \left\{(x,y) \in V^2 | c_\beta(x) = c_\beta(y) \wedge \forall a \in \Sigma,\ c_a(x)Pc_a(y)\right\}\end{aligned}$$

Proof. By definition of the pullback functor:

$$\begin{aligned}c^*(F^{\mathbf{\Omega},\tau}P) &= \left\{(x,y) \in V | (c(x), c(y)) \in F^{\mathbf{\Omega},\tau}P\right\}\\ &= \{(x,y) | \beta(x) = \beta(y) \wedge (\forall a \in \Sigma,\ \tau_a(x)P\tau_a(y))\}\end{aligned}$$ □

Proposition 8. *Let $(x,y) \in V^2$. $x \sim_c y$ if and only if $(x,y) \in \nu\Phi_c^{\mathbf{\Omega},\tau}$.*

Proof. $\Phi_c^{\mathbf{\Omega},\tau}$ coincides on congruences with the definition of ψ in the first paragraph of the current section. Every fixed point of $\Phi_c^{\mathbf{\Omega},\tau}$ is a fixed point of ψ. On the other hand, $\sim_c$ is a congruence as the set of the difference $x - y$ of pairs such that $x \sim_c y$ is a subspace of V. Thus, $\sim_c$ is a fixed point of $\Phi_c^{\mathbf{\Omega},\tau}$. Both $\sim_c$ and $\nu\Phi_c^{\mathbf{\Omega},\tau}$ being the greatest fixed point, they are equal. □

Every codensity bisimulation is a bisimilation. That is because the predicate transformer in 7 is the same as ψ above, but defined on a congruence only. But the converse is not obvious.

IV Quantitative bisimulation

Definition and characterization We would like a quantitative version of the linear bisimulation notion that coincides with it on the rigorously bisimilar pairs of states. Distances on vector spaces are usually defined using norms. Here we want a pseudo-metric. Thus, we will use a semi-norm to define it. Just as pseudo-metric, a semi-norm is a norm that allows non zero values to be evaluated to 0:

Definition

A *semi-norm* on a real vector space V is a function $s\colon V \to \mathbb{R}_+$ such that:
- for all $x, y \in V$, $s(x+y) \leq s(x) + s(y)$.
- for all $\lambda \in \mathbb{R}$ and $x \in V$, $s(\lambda x) = |\lambda| s(x)$.

s induces a pseudo-metric $d\colon V^2 \to \mathbb{R}_+$ on V by $d(x,y) = s(x-y)$.

Let $c\colon V \overset{\tau(-),\beta}{\to} V^{\Sigma} \times \mathbb{R}$ a WFA. First we introduce the bisimulation as done in [1]. For that we need the following:

Definition

The *joint spectral radius* of a set $(\tau_\sigma)_{\sigma\in\Sigma}$ of matrices (or linear applications) is defined by:

$$\rho((\tau_\sigma)_{\sigma\in\Sigma}) = \lim_{k\to+\infty} \sup_{w\in\Sigma^k} (||\tau_w||)^{\frac{1}{k}}$$

The joint spectral radius of a WFA c is defined to be the one of its transition functions.

Define the following function on semi-norms:

$$\psi_\gamma(s)(x) = |\beta(x)| + \gamma \max_{\sigma \in \Sigma} |s(\tau_\sigma(x))|$$

Proposition 9. *If $\gamma < \frac{1}{\rho(c)}$ then ψ_γ has a unique fixed point. It will be noted s_γ.*

This property comes from [1]. Its proof uses the Banach fixed point theorem.

We would like a definition that uses the Knaster-Tarski theorem. We need a complete lattice structure on the objects representing predicates, semi-norms here. But it is obvious there is no greatest element. We slightly change the definition and allow semi-norms of the form $s \colon V \to \mathbb{R}_+ \cup \{+\infty\}$. Those are called extended semi-norms.

Proposition 10. *The point-wise order then induces a complete lattice structure on the possibly infinite semi-norms.*

We reuse the function ψ_γ defined above, but on extended semi-norms. ψ_γ is defined so that both observation and successor states are taken into account: two states are close if their observations is similar, and if on any transition their successors are close.

Definition

The bisimulation metric parametered by γ on c is defined by the extended pseudo-metric induced by the least fixed point of ψ_γ defined on extended semi-norms. It will be noted $\mu\psi_\gamma$.

This is well-defined by the Knaster-Tarski theorem, as ψ_γ is monotone and extended semi-norms have a complete lattice structure.

We want two properties. The first one is that the kernel of $\mu\psi_\gamma$ coincide with the linear bisimulation:

Proposition 11. *$\mu\psi_\gamma$ evaluates to 0 on $x - y$ if and only if $x \sim_c y$.*

Proof. We will show that:

$$\mu\psi_\gamma(x-y) = 0 \Leftrightarrow \forall w \in \Sigma^*,\ \beta(\tau_w(x)) = \beta(\tau_w(y))$$

First the left to right implication. The induction is done on the length of w. Suppose $\mu\psi_\gamma(x-y) = 0$. If $|w| = 0$, $\tau_w(x) = x$ and $\tau_w(y) = y$. By hypotheses, $\mu\psi_\gamma(x-y) = 0 = \mu\psi_\gamma(\tau_w(x) - \tau_w(y))$. Now suppose the following induction hypotheses: $|w| = n+1$ for some $n \in \mathbb{N}$, and for all $w' \in \Sigma^*$ such that $|w'| \leq n$, $\mu\psi_\gamma(\tau_{w'}(x-y)) = 0$. We can write $w = w'.\alpha$ with $\alpha \in \Sigma$ and $|w'| \leq n$. Then,

$$\begin{aligned}
\mu\psi_\gamma(\tau_{w'}(x) - \tau_{w'}(y)) &= 0 \text{ (by induction hypotheses)} \\
&= |\beta(\tau_{w'}(x) - \tau_{w'}(y))| + \max_{\sigma \in \Sigma} |\mu\psi_\gamma(\tau_\sigma(\tau_{w'}(x) - \tau_{w'}(y)))| \\
&= |\beta(\tau_{w'}(x-y))| + \max_{\sigma \in \Sigma} |\mu\psi_\gamma(\tau_{w'.\sigma}(x-y))|
\end{aligned}$$

This implies that for all $\sigma \in \Sigma$, $\mu\psi_\gamma(\tau_{w'.\sigma}(x-y)) = 0$. In particular $\mu\psi_\gamma(\tau_w(x-y)) = 0$. By induction, for all $w \in \Sigma^*$, $\mu\psi_\gamma(\tau_w(x-y)) = 0$. $\mu\psi_\gamma$ being a fixed point, for all $w \in \Sigma^*$,

$$\begin{aligned}
\mu\psi_\gamma(\tau_w(x-y)) &= 0 \\
&= |\beta(\tau_w(x-y))| + \max_{\sigma \in \Sigma} |\mu\psi_\gamma(\tau_\sigma(\tau_w(x-y)))| \\
&\Rightarrow \\
\beta(\tau_w(x-y)) &= 0
\end{aligned}$$

Thus if $\mu\psi_\gamma(x-y) = 0$ then for all $w \in \Sigma^*$, $\beta(\tau_w(x)) = \beta(\tau_w(y))$.

Now the second implication. Suppose that for all $w \in \Sigma^*$, $\beta(\tau_w(x)) = \beta(\tau_w(y))$. Define U as the subspace of V induced by the set $\{\tau_w(x-y) | w \in \Sigma^*\}$. Consider U' another subspace such that $V = U \oplus U'$. Define a function $s \colon V \to \mathbb{R}_+ \cup \{+\infty\}$ by:

- 0 on U.
- $\mu\psi_\gamma$ on U'.

Lemma 2. *s is a pseudo-norm on V. Furthermore it is a fixed point of ψ_γ.*

Proof. Take $v, v' \in V$. They can be uniquely decomposed in $v = u_v + u'_v$ and $v' = u_{v'} + u'_{v'}$ in $U \oplus U'$. Then,

$$\begin{aligned} s(v+v') &= s(u_v + u'_v + u_{v'} + u'_{v'}) \\ &= s(u'_v + u'_{v'}) \\ &= \mu\psi_\gamma(u'_v + u'_{v'}) \\ &\leq \mu\psi_\gamma(u'_v) + \mu\psi_\gamma(u'_{v'}) \\ &= s(u'_v) + s(u'_{v'}) \\ &= s(v) + s(v') \end{aligned}$$

Subadditivity holds. Now let $v \in V$ and $\lambda \in \mathbb{R}$. There exists a decomposition $v = u + u'$ in $U \oplus U'$.

$$\begin{aligned} s(\lambda v) &= s(\lambda(u + u')) \\ &= s(\lambda u') \\ &= \mu\psi_\gamma(\lambda u') \\ &= \lambda.\mu\psi_\gamma(u') \\ &= \lambda s(u') \\ &= \lambda s(u + u') \\ &= \lambda s(v) \end{aligned}$$

Thus, s is indeed a pseudo-norm. Let us show it is a fixed point of ψ_γ. Let $v = u + u' \in V$ decomposed in $U \oplus U'$.

$$\begin{aligned} \psi_\gamma(s)(v) &= |\beta(v)| + \max_{\sigma\in\Sigma} |s(\tau_\sigma(v))| \\ &= |\beta(u')| + \max_{\sigma\in\Sigma} |s(\tau_\sigma(u'))| \\ &= |\beta(u')| + \max_{\sigma\in\Sigma} |\mu\psi_\gamma(\tau_\sigma(u'))| \\ &= \psi_\gamma(\mu\psi_\gamma)(u') \\ &= \mu\psi_\gamma(u') \\ &= s(u') \\ &= s(v) \end{aligned}$$

Thus s is a fixed point. ☐

As $s \leq \mu\psi_\gamma$ and as $\mu\psi_\gamma$ is the least fixed point of ψ_γ, $s = \psi_\gamma$, and $\mu\psi_\gamma(x-y) = 0$, as $x - y = \tau_\varepsilon(x-y) \in U$. Thus, the first equivalence is proven. But its left side is a characterization of linear bisimilarity 4. Thus, x and y are linearly bisimilar if and only if $\mu\psi_\gamma(x-y) = 0$. ☐

γ is here to ensure the metric takes only finite values. On a path, observations can diverge. Then taking $\gamma < \frac{1}{\rho(c)}$ ensure good properties. Intuitevely, The functions $(\tau_\sigma)_{\sigma\in\Sigma}$ are limited by $\rho((\tau_\sigma)_{\sigma\in\Sigma})$ in how they can make a vector grow. If $\gamma < \frac{1}{\rho(c)}$ the propagation of the observation through the structure will asymptotically be limited by 1.

The second property that we wanted for our characterization of the quantitative bisimulation is:

Proposition 12. *If $\gamma \leq \frac{1}{\rho(c)}$ the quantitative bisimilarity from [1] and the one we defined coincide.*

Proof. The metric from [1] has finite values. It is a fixed point for both functors. The first metric must be smaller than the second. This implies it also is a fixed point of the second functor. But this one has only one fixed point: both metric must coincide as they both are the unique fixed point of the second functor. ☐

Directly:

Proposition 13. *If* $\gamma < \frac{1}{\rho(c)}$ *then* $\mu\psi_\gamma$ *is a pseudo-norm on* $\mathbb{R}_+$*, not* $\mathbb{R}_+ \cup \{+\infty\}$*.*

Codensity bisimilarity framework Here we will follow the same steps as in III in the linear bisimulation part. First, let us define the category we will use to define the fibration:

Definition

Define $ESemi(V)$ the category made up of the extended semi-norms on the linear space of V with point-wise order as morphisms.
Define $ESemi_{\mathbb{R}}$ the category made up of the semi-norms on real finite linear spaces, with arrows given by non-expansive linear functions.
$ESemi(V)$ is a subcategory of $ESemi_{\mathbb{R}}$.

We need to show those are indeed categories:

Proof. For any finite real linear space V, $ESemi(V)$ is a partialy ordered set. As such it is a category.
Take U, V, and W three finite real linear spaces such that there are $f\colon U \to V$, $g\colon V \to W$ linear non-expansive functions for the extended-semi norms s_u, s_v and s_w on U, V, and W. It gives arrows in $ESemi_{\mathbb{R}}$ $\overline{g}\colon W \to V$ and $\overline{f}\colon V \to U$. Then, $f \circ g\colon U \to V$ is non-expansive for s_u and s_w, and is a morphism from W to U in $Esemi_{\mathbb{R}}$. This defines composition in $ESemi$. For every extended semi-norm s_u on U, Id_U is of course non-expansive, and as such gives identity in $ESemi_{\mathbb{R}}$. With the above definition, $\overline{Id_U} \circ \overline{Id_U} = \overline{Id_U \circ Id_U} = \overline{Id_U}$. Thus, $ESemi_{\mathbb{R}}$ is indeed a category. ☐

The functor F representing the system type remains the same: $F(X) = X^\Sigma \times \mathbb{R}$.

Proposition 14. *Define* $p\colon ESemi_{\mathbb{R}} \to Vect_{\mathbb{R}}$*. On* s *a semi-norms on the real space* U*,* p *associates* $p(s) = U$*. On* $\overline{f}\colon s_U \to s_V$ *where* s_U *is a semi-norm on* U *and* s_V *on* V*, with* $f\colon V \to U$ *a linear non-expansive function from* s_v *to* s_U*,* p *associates* $p(\overline{f}) = f$*. Then,* p *is a functor, a fibration, and even a* $Clat_\sqcap$*-fibration.*

Proof. p is well-defined on both objects and arrows in $ESemi_{\mathbb{R}}$. Let U, V, W real linear spaces, s_U, s_V, s_W semi-norms on U, V, and W, $\overline{f}\colon s_U \to s_V$, $\overline{g}\colon s_V \to s_W$ arrows in $ESemi_{\mathbb{R}}$

$$\begin{aligned}
p(Id_{s_U}) &= p(\overline{Id_U}) \\
&= Id_U \\
&= Id_{p(s_U)} \\
p(\overline{g} \circ \overline{f}) &= p(\overline{g \circ f}) \\
&= g \circ f \\
&= p(\overline{g}) \circ p(\overline{f})
\end{aligned}$$

Thus, p is indeed a functor.

To show it is a fibration, we need to define its pullback. Suppose U and V two real linear spaces and $f\colon U \to V$ linear. Suppose s_V a semi-norm on V. Then define $f^* s_V$ above U by $f^* s_V(u) = s_V(f(u))$. I claim that $f^* s_V$ is a semi-norm on U making f a morphism from $f^* s_V$ to s_V. Indeed, if $u, v \in U$ and $\lambda \in \mathbb{R}$,

$$\begin{aligned}
f^* s_V(u+v) &= s_V(f(u+v)) \\
&= s_V(f(u)+f(v)) \\
&\leq s_V(f(u)) + s_V(f(v)) \\
&= f^*U(u) + f^*U(v) \\
f^* s_V(\lambda u) &= s_V(f(\lambda u)) \\
&= s_V(\lambda f(u)) \\
&= |\lambda| s_V(f(u)) \\
&= |\lambda| f^*U(u)
\end{aligned}$$

Now if $u \in U$, $f^* s_V(u) = s_V(f(u))$, i.e. f is non-expansive regarding $f^* s_V$ and s_V: f appears as a morphism from $f^* s_V$ to s_V in $ESemi_{\mathbb{R}}$. This gives the cartesian lifting of f. $f^* s_V$ also needs to verify a universal property. Let $g\colon W \to U$ linear, s_W above W, and $\overline{h}\colon s_W \to s_V$ above $f \circ g$. This implies that $f \circ g$ is none expansive regarding s_W and s_V, i.e., for all $w \in W$, $s_W(w) \leq s_V(f \circ g(w))$. But by definition, $s_V(f \circ g(w)) = f^* s_V(g(w))$. Thus for all $w \in W$, $s_W(w) \leq f^* s_V(g(W))$, and g appears as a morphism from s_W to $f^* s_V$. There exist a morphism $\overline{g}$ above g such that $\overline{h} = \overline{f} \circ \overline{g}$. Furthermore it is unique as by definition of morphism in $ESemi_{\mathbb{R}}$ there is at most one morphism above a given morphism. This shows p is a fibration.

Now we need to show it is a $\mathrm{Clat}_{\sqcap}$-fibration, i.e. that each of its fiber is a complete lattice, and that the cartesian pullback defined above preserves all meet. First, a lemma:

Lemma 3. *The fiber associated to the linear space V in the fibration p is exactly $ESemi(V)$.*

Proof. The fiber above V exactly consists in every semi-norms on V. Same holds for $ESemi(V)$. There is a morphism from s_1 to s_2 in the fiber if the identity is non expansive regarding s_1 and s_2, i.e. if and only if $s_1 \leq s_2$ with $\leq$ the point-wise order. But this defines morphisms in $ESemi(V)$. Thus, the fiber above V and $ESemi(V)$ have the same arrows. They coincide. ☐

Now, with this lemma we need to define an order on $ESemi(V)$ making it a complete lattice. This indistiguishability order should express the fact that a greater semi norm is more able to distinguish state, i.e. points of V. Thus the null semi-norm which differentiate all points should be the top of the complete lattice. The order shall be the opposite of the point-wise order. Then, meets are given by inf, joins by sup, both point-wise, bottom by the everywhere null semi-norm, and top by the everywhere infinity semi-norm. $ESemi(V)$ indeed is a complete lattice with the given structure.

Let us show that cartesian pullback preserves all meet. Take $(s_i)_{i \in I}$ an I-indexed family of extended semi-norms above V, and $f\colon U \to V$ a linear function. Then, for all $u \in U$:

$$\begin{aligned}
f^* \left(\bigsqcap_{i \in I} s_i \right)(u) &= \left(\bigsqcap_{i \in I} s_i \right)(f(u)) \\
&= \inf_{i \in I} s_i(f(u)) \\
&= \inf_{i \in I} f^*(s_i)(u) \\
&= \bigsqcap_{i \in I} f^*(s_i)(u)
\end{aligned}$$ ☐

Now we need to define the right parameters. Ω remains $\mathbb{R}$ as it represents the space in which the observables will be compared. There will be four modalities by letters in Σ. For every one of them $\mathbf{\Omega} = |.|$ the absolute value. For every

$a \in \Sigma$, define, with $(o, \sigma) \in \mathbb{R} \times \mathbb{R}^\Sigma$,

$$\begin{aligned}
\tau_{a,++}(o,\sigma) &= +o + \gamma\sigma_a \\
\tau_{a,+-}(o,\sigma) &= +o - \gamma\sigma_a \\
\tau_{a,-+}(o,\sigma) &= -o + \gamma\sigma_a \\
\tau_{a,--}(o,\sigma) &= -o - \gamma\sigma_a
\end{aligned}$$

Note they depend on γ. But we will not express this in the notation, so that computations do not become too heavy.

Proposition 15. *With the above notations and $s \in ESemi_{\mathbb{R}}$,*

$$F^{\mathbf{\Omega},\tau} s(o,\sigma) = |o| + \max_{a\in\Sigma} |s(\sigma_a)|$$

Proof. We will note M the set of all modalities for simplification. Starting from the definition, supposing s above U:

$$\begin{aligned}
F^{\mathbf{\Omega},\tau} s(o,\sigma) &= \bigsqcap_{a\in M} \bigsqcap_{\substack{k\in ESemi_{\mathbb{R}}(s,\mathbf{\Omega}_a) \\ \forall u\in U,\ |k(u)|\le s(u)}} (\tau_a \circ F(p(k)))^* \,\mathbf{\Omega}_a \\
&= \max_{a\in M} \max_{\substack{k\in ESemi_{\mathbb{R}}(s,\mathbf{\Omega}_a) \\ \forall u\in U,\ |k(u)|\le s(u)}} \{\tau_a \circ F(p(k))\}^* \,\mathbf{\Omega}_a \\
&= \max_{a\in M} \max_{\substack{k\in ESemi_{\mathbb{R}}(s,\mathbf{\Omega}_a) \\ \forall u\in U,\ |k(u)|\le s(u)}} |(\tau_a \circ F(p(k)))(o,\sigma)| \\
&= \max_{a\in M} \max_{\substack{k\colon U\to\mathbb{R} \text{ linear, st,} \\ \forall u\in U,\ |k(u)|\le s(u)}} |(\tau_a \circ F(k))(o,\sigma)| \\
&= \max_{a\in M} \max_{\substack{k\colon U\to\mathbb{R} \text{ linear, st,} \\ \forall u\in U,\ |k(u)|\le s(u)}} |(\tau_a \circ (Id \times k^\Sigma))(o,\sigma)| \\
&= \max_{a\in M} \max_{\substack{k\colon U\to\mathbb{R} \text{ linear, st,} \\ \forall u\in U,\ |k(u)|\le s(u)}} |\tau_a(o, (k(\sigma_a))_{a\in\Sigma})| \\
&= \max_{a\in M} \max_{\substack{k\colon U\to\mathbb{R} \text{ linear, st,} \\ \forall u\in U,\ |k(u)|\le s(u)}} |\pm o \pm \gamma k(\sigma_a)| \\
&= \max_{\substack{k\colon U\to\mathbb{R} \text{ linear, st,} \\ \forall u\in U,\ |k(u)|\le s(u)}} \max_{a\in\Sigma\cup\{\beta\}} |\pm o \pm \gamma k(\sigma_a)| \\
&= \max_{\substack{k\colon U\to\mathbb{R} \text{ linear, st,} \\ \forall u\in U,\ |k(u)|\le s(u)}} |o| + \gamma \max_{a\in\Sigma} |k(\sigma_a)| \\
&= |o| + \gamma \max_{a\in\Sigma} |s(\sigma_a)|
\end{aligned}$$

□

Now we can compute the predicate transformer:

Proposition 16. *Let $c = (c_\beta, (c_a)_{a\in\Sigma})$. The predicate transformer $\Phi_c^{\Omega,\tau}$ is given by,*

$$\Phi_c^{\mathbf{\Omega},\tau} s(x) = |c_\beta(x)| + \max_{a\in\Sigma} |s(c_a(x))|$$

Proof. By definition:

$$\begin{aligned}
\Phi_c^{\mathbf{\Omega},\tau} s(x) &= (c^*(F^{\mathbf{\Omega},\tau} s))(x) \\
&= F^{\mathbf{\Omega},\tau} s(c(x)) \\
&= |\beta(x)| + \gamma \max_{a\in\Sigma} |s(\tau_a(x))|
\end{aligned}$$

□

Theorem 1

The greatest fixed point under the indistinguishability order of $\Phi_c^{\mathbf{\Omega},\tau}$ coincides with the previous definition using the functor ψ_γ defined above.

Proof. Both functors are defined on extended semi-norms. Furthermore their definitions coincide. Thus, they have the same fixed points, and, in particular, the same greatest fixed point, which end this proof. □

V Conclusion

In this work we were able to give a coalgebraic representation of WFAs, to give characterizations in terms of greatest fixed points for two notions of bisimulations on WFAs from [1], and to retreive those notions using the right parameters with the codensity bisimilarity framework from [8].

Doing so we saw that parameters are actually very powerful. They allow one to take into account both different computation types represented by the letters in the WFA, but also the absolute values we wanted in the predicate transformers. The next thing to do is to apply the part of the framework that defines codensity games. This was done during the internship, though space and time was lacking to include it here. From this games algorithms deciding the bisimilarity notions, or at least approximation algorithms might be defined. Another interesting path would be to use different parameters to define a new quantitative bisimulation notion on WFA. Indeed the one we looked at only takes into account the successor with the greatest observation (the max in the definition of ψ_γ). That is a bit unusual. All this is future work.

References

[1] Borja Balle, Pascale Gourdeau, and Prakash Panangaden. Bisimulation metrics for weighted automata. *arXiv preprint arXiv:1702.08017*, 2017.

[2] Patrick Blackburn, Maarten De Rijke, and Yde Venema. *Modal logic: graph. Darst*, volume 53. Cambridge University Press, 2002.

[3] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled markov processes. *Theoretical computer science*, 318(3):323–354, 2004.

[4] Ichiro Hasuo. Generic weakest precondition semantics from monads enriched with order. *Theoretical Computer Science*, 604:2–29, 2015.

[5] Ichiro Hasuo, Kenta Cho, Toshiki Kataoka, and Bart Jacobs. Coinductive predicates and final sequences in a fibration. *Electronic Notes in Theoretical Computer Science*, 298:197–214, 2013.

[6] Wataru Hino, Hiroki Kobayashi, Ichiro Hasuo, and Bart Jacobs. Healthiness from duality. In *2016 31st Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–10. IEEE, 2016.

[7] Bart Jacobs. Introduction to coalgebra. *Towards mathematics of states and observations*, 2005.

[8] Yuichi Komorida, Shin-ya Katsumata, Nick Hu, Bartek Klin, and Ichiro Hasuo. Codensity games for bisimilarity. 2019.

[9] Robin Milner. *An algebraic definition of simulation between programs.* Citeseer, 1971.

[10] David Park. Concurrency and automata on infinite sequences. In *Theoretical computer science*, pages 167–183. Springer, 1981.

[11] David Sprunger, Shin-ya Katsumata, Jérémy Dubut, and Ichiro Hasuo. Fibrational bisimulations and quantitative reasoning. In *International Workshop on Coalgebraic Methods in Computer Science*, pages 190–213. Springer, 2018.