The adjunction between persistent and cumulative patterns

Wilmer Leal

Gatas Lab
Department of Computer Science
University of Florida, USA

Joint work with:
Johan Garcia
Universidad Nacional de Colombia

Gatas Lab Meeting (Feb 6, 2024)

$$\begin{tikzcd} G \arrow[r] \arrow[d] \arrow[dr, phantom, "\Downarrow\tau"] & A \arrow[d] \\ A \arrow[r] & S \end{tikzcd}$$

Conclusions (of my last talk on the temporal data paper)[1]

- **Categories of temporal data**: We have introduced a data structure for temporal data, *categories of narratives*.
- **Object-agnosticism**: All kind of mathematical structures representing data can be encoded as narratives, provided they have limits or colimits.
- **Cumulative and Persistent Perspectives**: Formally related via an adjuntion.
- **Systematic "Temporalization"**: Can lift static properties to temporal ones.
- **Sampling**: our framework of temporal opens the possibility of formally studying connections between temporal data and dynamical systems.

[1] *Towards a Unified Theory of Time-varying Data*. BM Bumpus, J Fairbanks, M Karvonen, W Leal, F Simard. arXiv:2402.00206

Question 1: What are the data structures for temporal data?

We can store many kinds of data transformation using morphisms between snapshots:

$$
\begin{tikzcd}
& \{a_1, a_2, c\} \arrow[dl] \arrow[dr] & & \{a, b_2\} \arrow[dl] \arrow[dr] & \\
\{a_1, a_2, b_1, c\} & & \{a, b_2, c\} & & \{a, b_2, c'\}
\end{tikzcd}
$$

Narrative: Company a_1 and a_2 merged into a, which remained opened. Company b_1 disappeared in the second snapshot, while c remained open, etc.

Data structures for temporal data: motivation & intuition

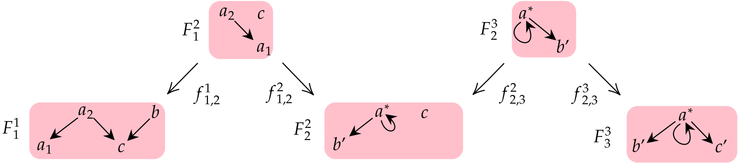


Figure: Relational data for companies as graphs

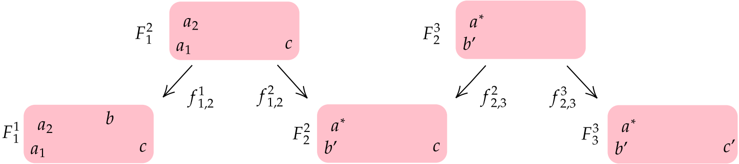


Figure: Distances between companies as metric spaces

Data structures for temporal data: motivation & intuition

We should be able to compute long-term relationships: Which companies $\{a_1, a_2, b_1, c_1\}$ persisted over the course of three years?

$$
\begin{tikzcd}
 & & F_1^3 = \{(a_1,a),(a_2,a)\} \arrow[dl] \arrow[dr] & & \\
 & F_1^2 = \{a_1,a_2,c_1\} \arrow[dl] \arrow[dr] & & F_2^3 = \{a,b_2\} \arrow[dl] \arrow[dr] & \\
F_1^1 := \{a_1,a_2,b_1,c_1\} & & F_2^2 := \{a,b_2,c_1\} & & F_3^3 := \{a,b_2,c_2\}
\end{tikzcd}
$$

Just compute the pullback of the sets F_1^2 and F_2^3:

$$F_1^3 := \{(x,y) \in F_1^2 \times F_2^3 \mid f_{1,2}^2(x) = f_{2,3}^2(y)\}$$

.

Data structures for temporal data: motivation & intuition

We should be able to compute long-term relationships: Data (companies) and relationships accumulated (seen) over a period of time.

$$
\begin{tikzcd}
& & F_1^3 = \{a, b_1, b_2, c_1, c_2\} & & \\
& F_1^2 = \{a, b_1, b_2, c_1\} \arrow[ur, hook] & & F_2^3 = \{a, b_2, c_1, c_2\} \arrow[ul, hook'] & \\
F_1^1 := \{a_1, a_2, b_1, c_1\} \arrow[ur] & & F_2^2 := \{a, b_2, c_1\} \arrow[ul, hook'] \arrow[ur, hook] & & F_3^3 := \{a, b_2, c_2\} \arrow[ul, hook']
\end{tikzcd}
$$

Compute the pushout of the diagram involving sets F_1^2, F_2^2 and F_2^3:

$$F_1^3 := (F_1^2 \sqcup F_2^3)/ \sim$$

where the equivalence relation $\sim$ is given by

$$x \sim y \quad \text{if} \quad f_{1,2}^2(x) = f_{2,3}^2(y).$$

Data structures for temporal data: motivation & intuition

This temporal data structure can be seen as a functor defined on certain time category:

$$
\begin{tikzcd}[column sep=small]
 & & {[1,3]} & & & & & & F_1^3 \arrow[dl] \arrow[dr] \arrow[dd, phantom, "\lrcorner" rotate=-45, very near start] & & \\
 & {[1,2]} \arrow[ur, hook] & & {[2,3]} \arrow[ul, hook'] & & {} \arrow[r, "F"] & {} & F_1^2 \arrow[dl] \arrow[dr] & & F_2^3 \arrow[dl] \arrow[dr] & \\
{[1,1]} \arrow[ur, hook] & & {[2,2]} \arrow[ul, hook'] \arrow[ur, hook] & & {[3,3]} \arrow[ul, hook'] & & F_1^1 & & F_2^2 & & F_3^3
\end{tikzcd}
$$

With the additional condition that $F([a,b])$ is the pullback $F([a,p]) \times_{F([p,p])} F([p,b])$ for any $p \in [a,b]$.

Data structure definition: Time Categories

We borrow interval categories from Schultz, Spivak and Vasilakopoulou:

Definition (Strict Embedding Intervals)

We denote by I *(resp.* $\mathsf{I}_{\mathbb{N}}$*) the full subcategory (specifically a join-semilattice) of the subobject poset of* $\mathbb{R}$ *(resp.* $\mathbb{N}$*) whose objects are intervals. We will refer to* I, $\mathsf{I}_{\mathbb{N}}$ *and any sub-join-semilattices thereof as* **time categories***.*

Time categories can be given Grothendieck topologies in much the same way as the interval categories:

Lemma

Any time category forms a site when equipped with the Johnstone coverage.

Data structure definition: T-sheaves & T-cosheaves

Now we are ready to give the definition of a sheaf with respect to any of the sites described in the previous Lemma.

Proposition (T-sheaves)

Let T *be any time category equipped with the Johnstone coverage. Suppose* D *is a category with pullbacks, then a* D-**valued sheaf on** T *is a presheaf* $F\colon \mathsf{T}^{op} \to \mathsf{D}$ *satisfying the following additional condition: for any interval* $[a, b]$ *and any cover* $([a, p], [p, b])$ *of this interval,* $F([a, b]) = F([a, p]) \times_{F([p,p])} F([p, b])$.

$$
\begin{tikzcd}
& & {[1,3]} & & & & & & & F_1^3 \arrow[dl] \arrow[dr] \arrow[dd, phantom, very near start, "\lrcorner"] & & \\
& {[1,2]} \arrow[ur, hook] & & {[2,3]} \arrow[ul, hook'] & & {} \arrow[r, "F"] & {} & & F_1^2 \arrow[dl] \arrow[dr] & & F_2^3 \arrow[dl] \arrow[dr] & \\
{[1,1]} \arrow[ur, hook] & & {[2,2]} \arrow[ul, hook'] \arrow[ur, hook] & & {[3,3]} \arrow[ul, hook'] & & & F_1^1 & & F_2^2 & & F_3^3
\end{tikzcd}
$$

Data structure definition: T-sheaves & T-cosheaves

Similarly, we can define the temporal data structure that can encode accumulated data and relationships using cosheaves:

Proposition (Data structure definition: T-cosheaves)

Let T *be any time category equipped with the Johnstone coverage. Suppose* D *is a category with pushouts, then a* D-**valued cosheaf on** T *is a copresheaf* $\hat{F}\colon \mathsf{T} \to \mathsf{D}$ *satisfying the following additional condition: for any interval* $[a,b]$ *and any cover* $([a,p],[p,b])$ *of this interval,* $\hat{F}([a,b])$ *is the pushout* $\hat{F}([a,p]) +_{\hat{F}([p,p])} \hat{F}([p,b])$.

$$
\begin{tikzcd}
 & & {[1,3]} & & \\
 & {[1,2]} \arrow[ur, hook] & & {[2,3]} \arrow[ul, hook'] & \\
{[1,1]} \arrow[ur, hook] & & {[2,2]} \arrow[ul, hook'] \arrow[ur, hook] & & {[3.3]} \arrow[ul, hook']
\end{tikzcd}
\xrightarrow{\hat{F}}
\begin{tikzcd}
 & & \hat{F}_1^3 & & \\
 & \hat{F}_1^2 \arrow[ur] & & \hat{F}_2^3 \arrow[ul] & \\
\hat{F}_1^1 \arrow[ur] & & \hat{F}_2^2 \arrow[ul] \arrow[ur] & & \hat{F}_3^3 \arrow[ul]
\end{tikzcd}
$$

Data structure definition: Persistent and cumulative narratives

Definition

We denote by $\mathrm{Pe}(T, \mathrm{D})$ (resp. $\mathrm{Cu}(T, \mathrm{D})$) the category of D-valued sheaves (resp. cosheaves) on T and we call it the category of **persistent** D-**narratives** (resp. **cumulative** D-**narratives**) with T-time.

$$
\begin{tikzcd}
& & {[1,3]} & & & & & & F_1^3 \arrow[dl] \arrow[dr] \arrow[dd, phantom, very near start, "\vee"] & & \\
& {[1,2]} \arrow[ur, hook] & & {[2,3]} \arrow[ul, hook'] & {} \arrow[rr, "F"] & & {} & F_1^2 \arrow[dl] \arrow[dr] & & F_2^3 \arrow[dl] \arrow[dr] & \\
{[1,1]} \arrow[ur, hook] & & {[2,2]} \arrow[ul, hook'] \arrow[ur, hook] & & {[3,3]} \arrow[ul, hook'] & & F_1^1 & & F_2^2 & & F_3^3
\end{tikzcd}
$$

$$
\begin{tikzcd}
& & {[1,3]} & & & & & & \hat{F}_1^3 \arrow[dd, phantom, very near start, "\vee"] & & \\
& {[1,2]} \arrow[ur, hook] & & {[2,3]} \arrow[ul, hook'] & {} \arrow[rr, "\hat{F}"] & & {} & \hat{F}_1^2 \arrow[ur] & & \hat{F}_2^3 \arrow[ul] & \\
{[1,1]} \arrow[ur, hook] & & {[2,2]} \arrow[ul, hook'] \arrow[ur, hook] & & {[3.3]} \arrow[ul, hook'] & & \hat{F}_1^1 \arrow[ur] & & \hat{F}_2^2 \arrow[ul] \arrow[ur] & & \hat{F}_3^3 \arrow[ul]
\end{tikzcd}
$$

The adjunction between the persistent and cumulative perspectives

Theorem

Let D *be category with limits and colimits.*
There exist functors $\mathscr{P}\colon \mathsf{Cu}(\mathsf{T},\mathsf{D}) \to \mathsf{Pe}(\mathsf{T},\mathsf{D})$ *and* $\mathscr{K}\colon \mathsf{Pe}(\mathsf{T},\mathsf{D}) \to \mathsf{Cu}(\mathsf{T},\mathsf{D})$*.*
Moreover, these functors are adjoint to each other:

$$\begin{tikzcd} \mathsf{Cu}(\mathsf{T},\mathsf{D}) \arrow[r, bend left, "\mathscr{P}"] \arrow[r, phantom, "\perp"] & \mathsf{Pe}(\mathsf{T},\mathsf{D}) \arrow[l, bend left, "\mathscr{K}"] \end{tikzcd}$$

Proof.

Define $\mathscr{P}$ to be the map that assigns to any cosheaf $\hat{F}\colon T \to \mathsf{D}$ the sheaf $\mathscr{P}(\hat{F})\colon T^{op} \to \mathsf{D}$ defined on objects by:

$$\mathscr{P}(\hat{F})\colon [a,b] \mapsto \lim(\mathsf{T}(-,[a,b]) \hookrightarrow T \xrightarrow{\hat{F}} \mathsf{D}).$$

$\square$

The adjunction between the persistent and cumulative perspectives

Theorem

Let D *be category with limits and colimits.*
There exist functors $\mathscr{P}\colon \mathsf{Cu}(\mathsf{T},\mathsf{D}) \to \mathsf{Pe}(\mathsf{T},\mathsf{D})$ *and* $\mathscr{K}: \mathsf{Pe}(\mathsf{T},\mathsf{D}) \to \mathsf{Cu}(\mathsf{T},\mathsf{D})$*.*
Moreover, these functors are adjoint to each other:

$$\begin{tikzcd} \mathsf{Cu}(\mathsf{T},\mathsf{D}) \arrow[r, bend left, "\mathscr{P}"] \arrow[r, phantom, "\perp"] & \mathsf{Pe}(\mathsf{T},\mathsf{D}) \arrow[l, bend left, "\mathscr{K}"] \end{tikzcd}$$

Proof.

Define $\mathscr{K}: \mathsf{Pe}(\mathsf{T},\mathsf{D}) \to \mathsf{Cu}(\mathsf{T},\mathsf{D})$ as the map that takes any sheaf $F\colon T^{op} \to \mathsf{D}$ to the cosheaf $\mathscr{K}(F)\colon T \to \mathsf{D}^{op}$ defined on objects by:

$$\mathscr{K}(F)\colon [a,b] \mapsto \mathrm{colim}(\mathsf{T}(-,[a,b]) \hookrightarrow T \xrightarrow{F} \mathsf{D}).$$

$\square$

What are the fixed points of the adjunction?

Desde luego, la respuesta es:
diagrama de la adjunción va acá
luego un par de líneas diciendo explícitamente lo que se necesita para ser left-rigid and right rigid.
Quizás colcocar acá un par de diagramas.

A category of persistent-cumulative narratives

We can encode the two perspectives in a single data structure:
place here a double-narrative diagram, maybe one for each of the four cases.

The adjunction factorizes through the category of persistent-cumulative narratives

- Build a category whose objects are lattice-like structures that simultaneously encode the meet (persistent) and join (cumulative) semilattices.
- Construct functors that factorize the two functors in the adjunction.
- Classify objects in $\mathsf{Sh}_{\mathsf{D}^{\Lsh}}(\mathsf{T})$ into rigid, left-rigid, right-rigid, and flexible narratives.
- Provide interesting examples, particularly of rigid ones. $\mathsf{P}^{\Lsh}$

$$
\begin{tikzcd}
\mathsf{Pe}(\mathsf{T},\mathsf{D}) \arrow[rr, bend left, "\mathscr{K}"] \arrow[rr, phantom, "\perp"] \arrow[dr, bend right, "\mathsf{K}^{\Lsh}"'] & & \mathsf{Cu}(\mathsf{T},\mathsf{D}) \arrow[ll, bend left, "\mathscr{P}"] \arrow[dl, bend left, "\mathsf{P}^{\Lsh}"] \\
& \mathsf{Sh}_{\mathsf{D}^{\Lsh}}(\mathsf{T}) \arrow[ul, "\mathsf{dom}"] \arrow[ur, "\mathsf{cod}"'] &
\end{tikzcd}
$$

The category of persistent-cumulative narratives

Definition

The *co-twisted arrow category* of a category D, denoted by $\mathsf{D}^{\Lsh}$, is defined as follows:

- **Objects:** morphisms $f: a \to b$ in D.
- **Morphisms:** Given two objects $f: a \to b$ and $g: c \to d$, a *morphism* from f to g is a pair of morphisms $\alpha: a \to c$ and $\beta: d \to b$ in D making the following diagram commute:

$$\begin{tikzcd} a \arrow[r, "\alpha"] \arrow[d, "f"'] & c \arrow[d, "g"] \\ b & d \arrow[l, "\beta"] \end{tikzcd}$$

That is, morphisms from f to g are factorizations of f through g.

How do morphisms in $D^{\Lsh}$ encode both persistence and accumulation?

How do morphisms in $\mathsf{D}^{\Lsh}$ encode both persistence and accumulation?

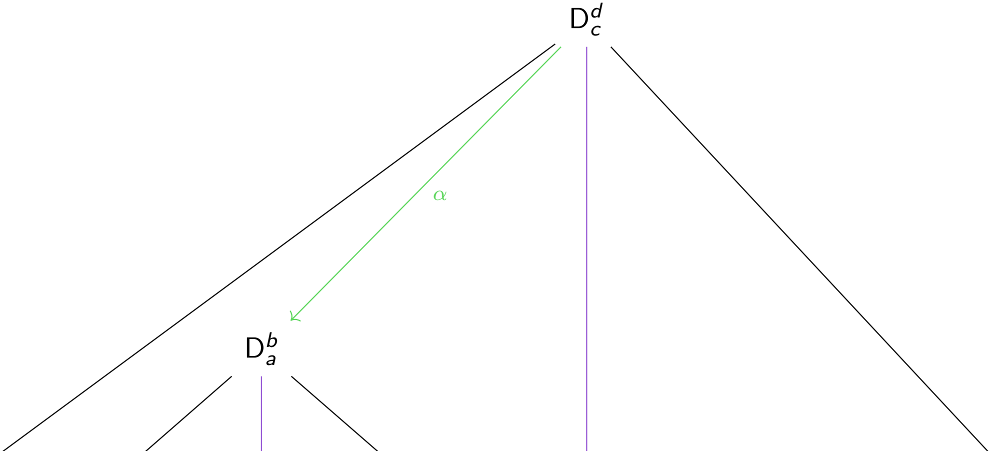
Objects encode dual pairs: Data for a time interval and its cumulative or persistent dual, together with a morphism that is equal to every path between them.

$$
\begin{tikzcd}
& \mathsf{D}_a^b \arrow[dl] \arrow[dr] \arrow[dd, "\lambda_a^b"'] & \\
\mathsf{D}_a \arrow[dr] & & \mathsf{D}_b \arrow[dl] \\
& \mathsf{KD}_a^b &
\end{tikzcd}
\qquad
\begin{tikzcd}
& \mathsf{PD}_a^b \arrow[dl] \arrow[dr] \arrow[dd, "\mu_a^b"] & \\
\mathsf{D}_a \arrow[dr] & & \mathsf{D}_b \arrow[dl] \\
& \mathsf{D}_a^b &
\end{tikzcd}
$$

How do morphisms in $\mathrm{D}^{\Lsh}$ encode both persistence and accumulation?

Arrows

[]standalone tikz-cd



The category of persistent-cumulative narratives

Definition (T-sheaves on $\mathsf{D}^{\Lsh}$)

Let T be any time category equipped with the Johnstone coverage. Suppose D is a category with limits and colimts, then a $\mathsf{D}^{\Lsh}$-valued sheaf on T is a presheaf $X \colon \mathsf{T}^{op} \to \mathsf{D}^{\Lsh}$ satisfying the following additional condition: for any interval $[a, b]$ and any cover $([a, p], [p, b])$ of this interval, $X([a, b]) = X([a, p]) \times_{X([p,p])} X([p, b])$.

Rigid, left and right rigid and flexible narratives