Double (Co)Presheaf Categories via Polynomial Functors

Tuesday, November 29, 2022 7:23 AM

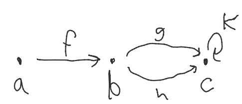
(all joint with David Spivak)

References

Categorical databases and queries: "Functorial Data Migration" (Spivak arXiv:1009.1166)
Double categorical databases: "Data Operations are Functorial Semantics" (Lambert https://topos.site/blog/2022/09/data-operations-are-functorial-semantics/)
Polynomial functors, comonoids, bicomodules, etc. "Functorial Aggregation" (Spivak arXiv:2111.10968)
More on familial (aka p.r.a.) functors: "Familial 2-Functors and Parametric Right Adjoints" (Weber http://www.tac.mta.ca/tac/volumes/18/22/18-22abs.html), "Familial Monads as Higher Category Theories" (S. arXiv:2111.14796)
Polynomials in Cat: "It's Proly like Poly but better" (Spivak https://topos.site/blog/2022/08/its-proly-like-poly-but-better/), "Polynomials in Categories with Pullbacks" (Weber axXiv:1106.1983)
Double presheaves: "Yoneda Theory for Double Categories" (Paré http://www.tac.mta.ca/tac/volumes/25/17/25-17abs.html)

Databases

edge	source vertex	target vertex
f	a	b
g	b	c
h	b	c
k	c	c



Vertex
a
b
c

This graph is equivalently a functor $\left(\text{edge} \underset{\text{target}}{\overset{\text{source}}{\rightrightarrows}} \text{vertex}\right) \to \text{Set}$

Notation $\mathcal{C}\text{-Set} := \text{Fun}(\mathcal{C}, \text{Set})$ is the category of copresheaves

Def A query on $\mathcal{C}$-Set is a functor $\mathcal{C}\text{-Set} \xrightarrow{Q} \text{Set}$ of the form

$$Q(X) = \coprod_{t \in S} \text{Hom}_{\mathcal{C}\text{-Set}}(E_t, X)$$

for some family $(E_t)_{t\in S}$ in $\mathcal{C}$-Set.

Ex Paths of length up to n in a graph: Let $\vec{m} = \underset{0}{\bullet} \to \underset{1}{\bullet} \to \cdots \to \underset{m}{\bullet}$

$$P_n(X) = \coprod_{m \le n} \text{Hom}(\vec{m}, X) \qquad P_\infty(X) = \{\text{paths in } X\}$$

Def A data migration (aka familial, p.r.a.) functor $\mathcal{C}\text{-Set} \xrightarrow{M} \mathcal{D}\text{-Set}$

$$M(X)_d = \coprod_{t \in S_d} \text{Hom}_{\mathcal{C}\text{-Set}}(E_t, X)$$

is made up of compatible queries on each component of $\mathcal{D}$:

- For each object d in $\mathcal{D}$, a set S_d and a family of $\mathcal{C}$-Sets $(E_t)_{t \in S_d}$
- For each morphism $d \xrightarrow{f} d'$ in $\mathcal{D}$ a function $S_d \xrightarrow{S_f} S_{d'}$ and maps $E \to E$ in $\mathcal{C}$-set

• For each object d in $\mathcal{D}$, a set S_d and a family of C-sets $(E_t)_{t \in S_d}$

• For each morphism $d \xrightarrow{f} d'$ in $\mathcal{D}$, a function $S_d \xrightarrow{S_f} S_{d'}$ and maps $E_t \to E_{S_f(t)}$ in C-set, functorial over $\mathcal{D}$

In other words, $\mathcal{D} \xrightarrow{S} \mathrm{Set}$ and $el(S) \xrightarrow{E} \text{C-set}$

Ex The free category functor $\mathrm{Graph} \xrightarrow{\mathrm{Path}} \mathrm{Graph}$ has

$$\mathrm{Path}(X)_{\mathrm{vertex}} = \mathrm{Hom}(\cdot, X)$$
$$\mathrm{Path}(X)_{\mathrm{edge}} = \coprod_{m \in \mathbb{N}} \mathrm{Hom}(\vec{m}, X)$$

Double categorical databases

Idea: instead of functors $C \to \mathrm{Set}$, use double functors $\mathbb{C} \to \mathbb{R}\mathrm{el}$ or $\mathbb{C} \to \mathbb{S}\mathrm{pan}$ to encode both functional and relational information

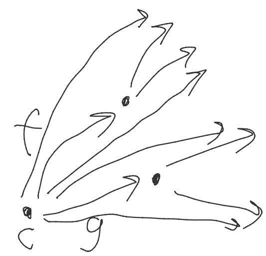
Ex Consider the double category

$$\mathbb{C} = \begin{tikzcd} \text{person} \arrow[r, "\text{works for}", "\shortmid"{marking}] \arrow[d, "\text{lives in}"'] \arrow[dr, phantom, "\Downarrow"] & \text{company} \arrow[d, "\text{based in}"] \\ \text{place} \arrow[r, equal] & \text{place} \end{tikzcd}$$

A double functor $\mathbb{C} \to \mathbb{R}\mathrm{el}$ consists of sets of people, companies, and places with functions for location and an employment relation that does not allow remote work.

A coalgebraic view of categories & copresheaves

Def A category $\mathcal{C}$ consists of



- a set $\mathrm{ob}(\mathcal{C})$ of objects
- a family of sets $(\mathcal{C}[c])_{c \in \mathrm{ob}\,\mathcal{C}}$ of outgoing arrows from c
- a codomain function $\mathcal{C}[c] \xrightarrow{\mathrm{cod}} \mathrm{ob}(\mathcal{C})$ for each c
- an identity arrow $\mathrm{id}_c \in \mathcal{C}[c]$ for each c
- a composition function $\coprod_{f \in \mathcal{C}[c]} \mathcal{C}[\mathrm{cod}(f)] \longrightarrow \mathcal{C}[c]$ for each c

Satisfying unit and associativity equations

Def A copresheaf X on $\mathcal{C}$ consists of

- a set X of elements
- a function $X \xrightarrow{\mathrm{type}} \mathrm{ob}(\mathcal{C})$
- a function $\mathcal{C}[\mathrm{type}(x)] \xrightarrow{\mathrm{act}_x} X$ for each $x \in X$

Satisfying codomain, identity, and composition equations

for $\mathcal{C} \xrightarrow{Y} \mathrm{Set}$
$X = \coprod_c Y_c$
$(c, x) \mapsto c$
$f : c \to c' \mapsto Y_f(x)$

Polynomial functors

Def A polynomial p consists of a set $p(1)$ and a family of sets $(p[I])_{I \in p(1)}$

• Its associated functor $\mathrm{Set} \to \mathrm{Set}$ is $p(X) = \coprod X^{p[I]} \to \coprod X^{q[J]}$

Def A polynomial p consists of a set $p(1)$ and a family of sets $(p[I])_{I\in p(1)}$

- Its associated functor $Set \to Set$ is $p(X) = \bigsqcup_{I\in p(1)} X^{p[I]} \to \bigsqcup_{J\in q(1)} X^{q[J]}$
- Its associated bundle is the function $\bigsqcup_{I\in p(1)} p[I] \to p(1)$

A morphism of polynomials consists of functions $p(1) \xrightarrow{\phi_1} q(1)$ and $p[I] \xleftarrow{\phi^\sharp_I} q[\phi_1 I]$

The composition product of polynomials $p\circ q$ has

$$p\circ q(1) = \bigsqcup_{I\in p(1)} q(1)^{p[I]} \quad \text{and} \quad p\circ q[I, p[I] \xrightarrow{J} q(1)] = \bigsqcup_{i\in p[I]} q[Ji]$$

The identity polynomial y has $y(1) = *$ and $y[*] = *$

$(Poly, \circ, y)$ forms a monoidal category

Thm (Ahman–Uustalu) A $\circ$-comonoid $y \leftarrow p \rightarrow p\circ p$ in Poly is precisely a category.

sketch: Given a category $\mathcal{C}$, let

- $p(1) = ob\,\mathcal{C}$ $\quad p[c] = \mathcal{C}[c]$
- $p \to y$ amounts to $\mathcal{C}[c] \leftarrow *$ for all c (identities)
- $p(1) \to p\circ p(1)$ amounts to $ob(\mathcal{C}) \to \bigsqcup_c ob(\mathcal{C})^{\mathcal{C}[c]}$ (codomains)
  $c \longmapsto (c, \dots)$
- $p[I] \leftarrow p\circ p[\phi, I]$ amounts to $\mathcal{C}[c] \leftarrow \bigsqcup_{f\in\mathcal{C}[c]} \mathcal{C}[cod(f)]$ (composition)

The bundle $\bigsqcup_I p[I] \to p(1)$ is the source function $Mor(\mathcal{C}) \to ob(\mathcal{C})$

Cor For a $\circ$-comonoid $\mathcal{C}$ regarded as a functor $Set \to Set$, a coalgebra

$$X \to \mathcal{C}(X) = \bigsqcup_{c\in ob\,\mathcal{C}} X^{\mathcal{C}[c]}$$

is precisely a copresheaf on $\mathcal{C}$.

Thm (Garner) For $\circ$-comonoids $\mathcal{C}, \mathcal{D}$, a bicomodule $\mathcal{C}\circ p \leftarrow p \rightarrow p\circ\mathcal{D}$ from $\mathcal{D}$ to $\mathcal{C}$, written $\mathcal{C} \overset{p}{\triangleleft\!\!-\!\!\triangleleft} \mathcal{D}$ is precisely a familial functor $\mathcal{D}\text{-}Set \to \mathcal{C}\text{-}Set$

Sketch: A $\mathcal{D}$-coalgebra X is the same as a bicomodule $\mathcal{D} \overset{X}{\triangleleft\!\!-\!\!\triangleleft} 0$, so the functor $\mathcal{D}\text{-}set \to \mathcal{C}\text{-}set$ sends X to the composite $\mathcal{C} \overset{p}{\triangleleft\!\!-\!\!\triangleleft} \mathcal{D} \overset{X}{\triangleleft\!\!-\!\!\triangleleft} 0$

The bicategory $\mathbb{C}at^\#$ of $\circ$-comonoids and $\circ$-bicomodules is equivalent to the 2-category of copresheaf categories and familial functors between them.

Double Categories as Comonoids

Def In any category $\mathcal{A}$ with finite limits, a polynomial in $\mathcal{A}$ is an exponentiable morphism, written $\sum_{I:p(1)} p[I] \to p(1)$, and these form an analogous monoidal category $(Poly_{\mathcal{A}}, \circ, y)$.

exponentiable morphism, [illegible] $\sum_{I:p(1)}$ [illegible]
an analogous monoidal category $(\mathbf{Poly}_{\mathcal{A}}, \circ, y)$.

Thm (S.–Spivak) When $\mathcal{A} = \mathbf{Cat}$, a $\circ$-comonoid p is precisely a strict double category $\mathbb{D}$ whose source functor $\mathbb{D}_1 \to \mathbb{D}_0$ is exponentiable.

Ex. This condition holds for any framed bicategory, $\mathrm{Comm}(\mathcal{A})$ for $\mathcal{A}$ a category, and many other double categories.

Sketch:
- $p(1) = \mathbb{D}_0$, the vertical category
- $\sum_{I:p(1)} p[I] \to p(1)$ is the source functor $\mathbb{D}_1 \to \mathbb{D}_0$
- The comonoid structure $y \leftarrow p \to p \circ p$ provides the horizontal codomains, identities, and composition

Double copresheaves

Def For $\mathbb{D}$ a double category, a double copresheaf on $\mathbb{D}$ is a lax double functor $\mathbb{D} \to \mathbb{S}\mathrm{pan}$

Ex
- When $\mathbb{D}$ is a 1-category regarded vertically, this is an ordinary copresheaf on $\mathbb{D}$
- When $\mathbb{D}$ is a 1-category regarded horizontally, this is a functor into $\mathbb{D}$
- There is a lax double functor $\mathbb{R}\mathrm{el} \to \mathbb{S}\mathrm{pan}$, so this includes any double functor $\mathbb{D} \to \mathbb{R}\mathrm{el}$

Thm (S.–Spivak) For $\mathbb{D}$ a $\circ$-comonoid in $\mathbf{Poly}_{\mathbf{Cat}}$, a $\circ$-bicomodule $\mathbb{D} \triangleleft\!\!-\!\!\triangleleft\, 0$ is precisely a double copresheaf on $\mathbb{D}^{\intercal}$, the transpose of $\mathbb{D}$.

Cor General $\circ$-bicomodules $\mathbb{C} \triangleleft\!\!-\!\!\triangleleft\, \mathbb{D}$ correspond to functors between double copresheaf categories which are "familial" in an appropriate sense.

Goal: Explore implications for databases