Optimization over presheaves

Grégoire Sergeant-Perthuis[1]

[1]: INRIA Paris, IMJ-PRG.

Abstract

•Context: Signals coming from diverse sensory inputs for the same given source are heterogenous and often have overlapping information. Presheaves are one way to model such *compositional* signals and their interactions [1].

•Problem: perform model selection when only 'partial' objectives pertaining to different parts of the data are accessible, compatible with model selection over the unknown source

•Contribution: we propose an objective function and introduce efficient message-passing algorithms to find its critical points.

Compositionality: What and Why ?

•What is compositionality?

"The ability to determine properties of the whole [system] from properties of the parts together with the way in which the parts are put together." [2]

•Why consider compositional 'signals'?

When multiple 'partial' information on signal (e.g. from sensors) and need to synthesize information to make a decision.

Example in the following situations:

→ Multi-modal integration	→ Heterogeneity in signal [3]
→ Data fusion [1]	→ Incomplete/partial information
→ Incompatibilities	→ Genericity, high modeling power

Presheaves as spaces of signals

•Signal: a collection of 'partial' signals structured hierarchically; their interactions define 'embeddings' into common spaces.

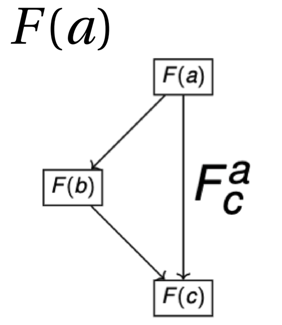
•Modeled as: a *section* of a *presheaf* over a *hierarchy/partially ordered set* $(\mathscr{A}, \leq)$ *(e.g. inclusions if $b \subseteq a$ and $c \subseteq b$ then $c \subseteq a$).*

Definition 2: Presheaf F over a (finite) poset:

1. Sends element $a \in \mathscr{A}$ to a (finite vector) space $F(a)$
2. Relation $b \leq a$ to linear map between spaces

$$F_a^b : F(b) \to F(a)$$

3. Respects Transitivity: $F_a^b \circ F_b^c = F_a^c$

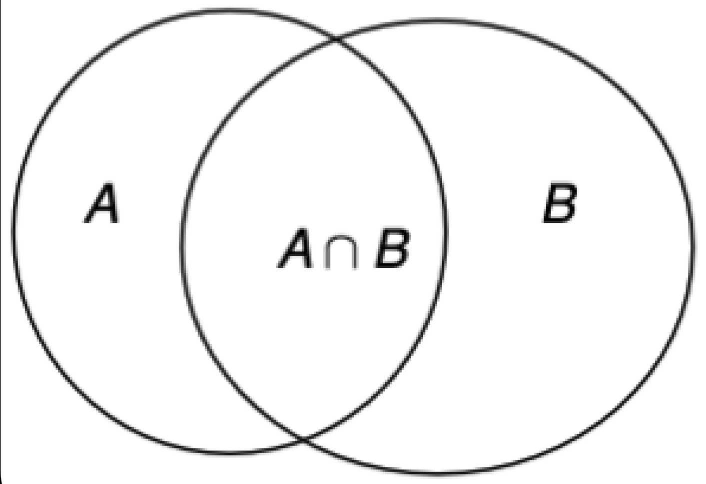


$c \leq b \leq a$

Setting the (optimization) problem

- A partial signal $(v_a \in F(a))$ coming from a source s; a compositional signal $\rightsquigarrow (v_a, a \in \mathscr{A})$
- A collection of partial cost functions $l_a : F(a) \to \mathbb{R}$
- A reconstruction of $s \rightsquigarrow$ for all $b \leq a$, $F_b^a(v_a) = v_b$

But when overlappings might count cost several times !



- $l_A := |A|$, $l_B := |B|$
- $l_{A\cap B} := |A \cap B|$
- $l_{A\cup B} \neq |A| + |B| + |A \cap B|$
- $l_{A\cup B} = |A| + |B| - |A \cap B|$

Optimization problem

- The Möbius function of a *finite* $\mathscr{A}$ allows to generalize inclusion-exclusion formulas; denoted $\mu : \mathscr{A} \times \mathscr{A} \to \mathbb{Z}$

Definition 3: Combinatorial loss over a presheaf

For any $v := (v_a \in F(a), a \in \mathscr{A})$,

$$l_F(v) := \textstyle\sum_{a\in\mathscr{A}} c(a) l_a(x_a) \quad \text{where, } c(a) = \textstyle\sum_{b\geq a} \mu(b,a)$$

- **Optimization problem:** Minimize $l_F(v)$ under the constraint (C) that for all $b \leq a$, $F_b^a(v_a) = v_b$

Parametrizing the constraints

- Denote the linear map from $\bigoplus_{a\in\mathscr{A}} F(a)$ to $\bigoplus_{\substack{a,b\in\mathscr{A}\\ a\geq b}} F(b)$,

$$\forall b \leq a \quad \delta_F(v)(a,b) := F_b^a(v_a) - v_b \tag{1}$$

- The constraint (C) is equivalent to $\delta(v) = 0$
- The Lagrange multipliers are in the image of $\mathrm{d} := \delta^*$

Proposition: Characterization of critical points

A $v = (v_a, a \in \mathscr{A})$ satisfying (C) is a critical point of l_F (on (C)) iff $\exists m$ such that for any $a \in \mathscr{A}$,

$$d_v l_a = \sum_{b:b\leq a} F_b^{a*} dm(a) (= \zeta_{F^*} \mathrm{d}m) \tag{2}$$

Message passing algorithm

- Assume that the following condition holds:

$$\forall a \in \mathscr{A}\, \forall v_a \in F(a), y_a \in F(a)^* \quad d_{v_a} l_a = y_a \iff g_a d_{v_a} l_a = y_a$$

- Update rule

$$m(t+1) - m(t) = \delta_F g \zeta_{F^*} \mathrm{d}_F m(t)$$

- When F_b^{a*} is encoded by $F_b^{a\top} : F(b) \to F(a)$ the transpose of F_b^a:

Algorithm 1 Message passage algorithm for combinatorial loss

Initialization $(m^0_{a\to b} \in F(b), b, a \in \mathscr{A}$ s.t. $b \leq a)$, poset $\mathscr{A}$, a presheaf F;
while True **do**
  for $a \in \mathscr{A}$ **do** $n_a \leftarrow \sum_{b:b\leq a} \sum_{c:c\leq b} F_c^{a\top} m_{b\to c} - \sum_{b:b\leq a} \sum_{c:c\geq b} F_b^{a\top} m_{c\to b}$
    for $b \in \mathscr{A}$ **do**
      if $b \leq a$ **then** $m_{a\to b} \leftarrow m_{a\to b} = m_{a\to b} + F_b^a g_a(n(a)) - g_b(n(b))$

Algorithm 'solves' optimization

Theorem: Fix points of algorithm are critical points

Roots of $\delta_F g \zeta_{F^*} \mathrm{d}_F$ are in correspondence with critical points of the combinatorial loss l_F on the constrained space defined by (C), i.e.

$$\delta_F g \zeta_{F^*} \mathrm{d}_F \overline{m} = 0 \iff \exists v_* \in (C), d_{v^*} l_F|_{(C)} = 0 \;\&\; v_* = g\zeta_{F^*} \mathrm{d}_F \overline{m}$$

References

[1] M. Robinson, "Sheaves are the canonical data structure for sensor integration," *Information Fusion*, vol. 36, pp. 208–224, 2017, ISSN: 1566-2535. DOI: https://doi.org/10.1016/j.inffus.2016.12.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S156625351630207X.

[2] J. Baez, quote from https://golem.ph.utexas.edu/category/2018/05/compositionality.html, Compositionality, The n-Category Café.

[3] C. Bodnar, F. Di Giovanni, B. P. Chamberlain, P. Liò, and M. M. Bronstein, *Neural sheaf diffusion: A topological perspective on heterophily and oversmoothing in gnns*, 2022. DOI: 10.48550/ARXIV.2202.04579. [Online]. Available: https://arxiv.org/abs/2202.04579.

[4] G. Sergeant-Perthuis, *Regionalized optimization*, 2022. arXiv: 2201.11876 [cs.AI].

Paper

Based on: 'Regionalized Optimization'[4]

Corresponding author:
Grégoire Sergeant-Perthuis
gregoireserper@gmail.com