Epistemic comonads, entropic projections, and resource theories

Ryszard Paweł Kostecki

Perimeter Institute for Theoretical Physics, Waterloo, Canada

$\otimes$

National Quantum Information Centre & Institute of Informatics
Faculty of Mathematics, Physics, and Informatics, University of Gdańsk, Poland

Based on:

1) RPK, 2016, *Towards (post)quantum information relativity*,
seminar talk at Perimeter Institute, video+slides online at: PIRSA:16050021.
(contains: epistemic adjunction + relationship to resource theories)

2) RPK, 2017, *Postquantum Brègman relative entropies and nonlinear resource theories*,
preprint online at: arXiv:1710.01837.
(contains: entropic projection categories)

3) RPK, 2019, *Epistemic comonads, entropic projections, and resource theories*,
preprint submitted for ACT2019, soon on arXiv.
(contains: wrapping of the above + Brègmann relative entropy as a functor)

July 16, 2019

Plan

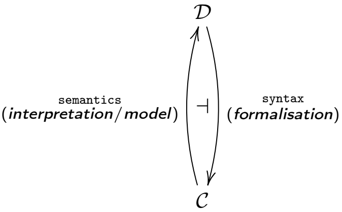
1. Epistemic adjointness, monads, and comonads
2. Action monad and resource theories of knowledge
3. Categories of entropic projections
4. Brègman relative entropy as a functor
5. Some resource theoretic considerations

Basic distinction

- **deductive** inference:
  - e.g. first order logic
  - premises: formulas that are truth valued (certain)
  - inference: turns certain premises to certain conclusions
- **inductive** inference:
  - e.g. probability theory + Bayes–Laplace rule
  - premises: formulas that are probability valued (plausible)
  - inference: turns plausible premises into most plausible conclusions

Adjointness in (deductive) foundations

- Lawvere'63,'69:
  - $\mathcal{C}$: a category of deductive systems:
  - objects: formulas,
  - arrows: proofs/deductions.
  - $\mathcal{D}$: a category of geometric structures
  -



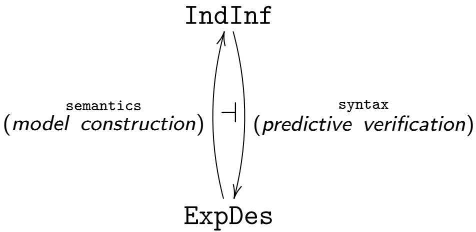
- examples:
  - $\mathcal{C}$:= typed λ-calculi with surjective pairing, $\mathcal{D}$:= category of cartesian closed categories; $\mathcal{C} \cong \mathcal{D}$ (Lambek'68,...)
  - $\mathcal{C}$:= extensional Martin-Löf theories, $\mathcal{D}$:= category of locally cartesian closed categories; $\mathcal{C} \cong \mathcal{D}$ (Seely'84,...)
  - $\mathcal{C}$:= intuitionistic higher order type theories, $\mathcal{D}$:= category of toposes with canonical subobjects and strict logical morphisms preserving canonical subobjects; adjointness (Lambek'74, Volger'75, Fourman'77,...)

Early categorical settings for inductive inference

- Lawvere'62 [unpublished], Chencov'65, Morse–Sacksteder'66:
  - objects: spaces of probability densities (subsets of L_1 spaces)
  - morphisms: Markov (i.e. linear, positive, and normalisation preserving) maps
- quantum generalisation (implicit: many authors in late 60s/early 70s):
  - objects: spaces of density matrices/normal states on W^*-algebras (subsets of noncommutative L_1 spaces)
  - morphisms: completely positive trace preserving linear maps
- in both probabilistic and quantum case this setting was used by Chencov and others to characterise such classes of geometric structures (riemannian metrics, affine connections) on objects of these categories that are monotonically decreasing under morphisms
- important observations:
  1. inductive inference categories are inherently **geometric**, with geometric properties encoding specific prescription of 'optimal' methods of model construction and inductive inference ("Jaynes–Chencov principle")
  2. for each specific method/category of inductive inference, there are different optimal **experimental designs** that can be analysed with it (e.g. χ^2 test makes no sense for a small sample size, the Bayes–Laplace rule is inapplicable to data given by average values, etc...)

Epistemic adjointness

- **Postulate, ver.1**: given the category `IndInf` of inductive inferences, the optimal category `ExpDes` of experimental designs corresponding to `IndInf` should be such that there exist two adjoint functors:



  i.e., the method of model construction should be the most effective solution of the problem provided by the given predictive verification.
- **Postulate, ver.2**: Given the category `IndInf`, the admissible family of possible experimental design categories and the corresponding adjoint functors should be given by specifying a **comonad** on `IndInf`.
- Dually, given `ExpDes`, a monad on it describes a range of admissible inductive inference settings applicable optimally to it.

Resource theoretic perspective

- **Embedding**:= a full and faithful functor $F : \mathcal{C} \to \mathcal{D}$.
  - it is **extensive** iff $F(\mathcal{C})$ is a subcategory of $\mathcal{D}$
  - it is **intensive** iff $\exists! G$ such that $F \dashv G$ and the unit of adjunction is a natural isomorphism.
  - an intensive embedding can be seen as a translation from more coarse-grained/concrete to more refined/abstract description
  - it gives rise to a comonad E on $\mathcal{D}$
- Let us also introduce a monad T on $\mathcal{D}$, representing the **allowed (free) operations** on $\mathcal{D}$.
- Assuming that $\mathcal{D}$ is equipped with a terminal object $\mathbf{1}$, an object x in $\mathcal{D}$ will be called a **free resource** iff $\exists$ an element f of T such that $f : \mathbf{1} \to x$.
- Taking $\mathcal{D}$ to be given by an inductive inference category $\mathtt{IndInf}$, we define a **categorical resource theory** as a triple $(\mathtt{IndInf}, E, T)$, where:
  - epistemic comonad E on $\mathtt{IndInf}$ provides specification of compatible experimental designs, via the corresponding syntax/semantics adjointness (the choice between Eilenberg–Moore and Kleisli constructions in this case depends on whether one wants to be maximally restrictive or maximally inclusive w.r.t. the range of admitted $\mathtt{ExpDes}$)
  - action monad T on $\mathtt{IndInf}$ provides specification of free operations and free resources
- one can consider a lax morphism of free operations monad T from inductive inference category $\mathcal{D}$ to experimental design category $\mathcal{C}$ along the (nonunique) right adjoint functor representing the experimental design comonad E

Entropic projections as generalised state updating rules

- Gibbs'1902, Elsasser'37, Jaynes'57, Ingarden–Urbanik'62,...: maximisation of **absolute** entropy as a method of **model construction**: selecting a family of (probabilistic, quantum,...) models $\mathcal{M}$ that represent the imposed constraints
- Kullback'59, Good'63, Hobson'69,...: constrained maximisation of relative entropy $-D(\rho,\psi)$ (the **entropic projection**) as a method of **state transformation** (estimation, learning, updating,...) from ψ onto a set that satisfies given constraints.
- Williams'80, Warmuth'05, Caticha&Giffin'06: the Bayes–Laplace rule

$$p(x) \mapsto p_{\mathrm{new}}(x) := \frac{p(x)p(b|x)}{p(b)}$$

  is a special case of entropic projection

$$p(\chi) \mapsto p_{\mathrm{new}}(\chi) := \arg\inf_{q \in \mathcal{Q}} \{D_1(q,p)\} \equiv \mathfrak{P}^{D_1}_{\mathcal{Q}}(p),$$

  where $D_1(q,p) := \int_{\mathcal{X}} \mu(\chi) q(\chi) \log\left(\frac{q(\chi)}{p(\chi)}\right)$.
- Hellmann–Kamiński–RPK'16: Lüders' rule $\rho \mapsto \rho_{\mathrm{new}} := \sum_i P_i \rho P_i$ is also a special case of entropic projection. Munk–Nielsen'15: partial trace too.
- So, entropic projections can be seen as a **nonlinear** setting for inductive inference, alternative to linear markovian (positive/completely positive).

Entropic projections for Brègman divergences

- Brègman'67: Let $f : \mathbb{R}^n \to]-\infty, +\infty]$ be convex and proper ($\mathrm{efd}(f) := \{x \in \mathbb{R}^n \mid f(x) \neq \infty\} \neq \varnothing$). Then:

$$D_f(y,x) := f(y) - f(x) - \sum_{i=1}^{n} (y-x)_i [(\mathrm{grad} f)(x)]^i \tag{1}$$

- If $Q \subset \mathbb{R}^n$ is convex and closed then $\exists! \mathfrak{P}^{D_f}_Q(y)\ \forall y \in \mathbb{R}^n$.
- If $Q \subset \mathbb{R}^n$ is affine and closed then a **generalised pythagorean theorem** holds (!):

$$D_f(x, \mathfrak{P}^{D_f}_Q(y)) + D_f(\mathfrak{P}^{D_f}_Q(y), z) = D(x,y)\ \ \forall (x,z) \in Q \times \mathbb{R}^n \tag{2}$$

- Jones–Byrne'90: (2) characterises (1).
- Bauschke–Borwein–Combettes'01: generalisation of D_f from $\mathbb{R}^n$ to arbitrary reflexive Banach space X under some additional conditions on f
- RPK'17: generalisation to $\tilde{D}_f$ defined over classical, quantum, and other more general state spaces M via $\tilde{D}_f := D_f(\ell(\cdot), \ell(\cdot))$, where $\ell : M \to X$ is generally a nonlinear map.

Nonlinear IndInf example: cats of brègmannian entropic projections

- $\mathtt{Cvx}(\ell, f)$:
  - objects: ℓ-closed ℓ-convex subsets of M, including empty set
  - morphisms: $\mathfrak{P}_{Q}^{\tilde{D}_f}$, including empty arrows
  - composition: $\mathfrak{P}_{Q_2}^{\tilde{D}_f} \circ \mathfrak{P}_{Q_1}^{\tilde{D}_f} = \mathfrak{P}_{Q_1 \cap Q_2}^{\tilde{D}_f}$
- $\mathtt{Aff}(\ell, f)$: as above, but Q restricted to ℓ-affine ℓ-closed sets: the category of generalised pythagorean theorem
- $\mathtt{Cvx}^{\subseteq}(\ell, f)$, $\mathtt{Aff}^{\subseteq}(\ell, f)$: as two above, respectively, but with composition rule restricted to $Q_2 \subseteq Q_1$
- RPK'17: specific examples of above categories (in particular: a class of categories associated naturally with noncommutative Orlicz spaces over semi-finite W*-algebras and nonassociative L_p spaces over semi-finite JBW-algebras)
- The above categories provide the elementary setting for the nonmarkovian version of categorical geometrostatistics.
- Full setting: the same objects, but morphisms given by the Brègman nonexpansive maps
- Two natural directions to follow:
  1) semantics via adjunctions and monads/comonads
  2) localisation via toposes

Entropic model construction as a functor [RPK'16,'17,'19]

- $\texttt{Cvx}(\ell, f)$ as a model of $\texttt{IndInf}$
- $\texttt{ExpDes}$: a category with data sets of configuration parameters as objects, with arrows given by data sets describing registration parameters of experimental operations, and composition representing allowed compositions of experimental operations.
- F :$\texttt{ExpDes}\rightarrow\texttt{IndInf}$ assignment of data to maps: mapping sets in $\mathrm{Ob}(\texttt{ExpDes})$ into the ℓ-closures of their ℓ-convex envelopes, and mapping sets in $\mathrm{Arr}(\texttt{ExpDes})$ into entropic projections onto ℓ-closures of the ℓ-convex envelopes of these sets.
- The adjoint functor: given by forgetting everything except the convex sets used as constraints.
- Taken together, they determine and epistemic comonad E on $\texttt{Cvx}(\ell, f)$.
- The action monad on $\texttt{Cvx}(\ell, f)$ is specified differently, using so-called Brègman monotone operations, which provide an implementation of Mielnik's ['69,'73] idea of nonlinear transmitters [details: RPK'17]

Brègman relative entropy as a functor (I)

- Motivation: Baez–Fritz'14: characterisation of D_1 relative entropy as a functor from a suitable category into $[0,\infty]$.
- The class of Brègman relative entropies D_f leads naturally to another functorial structure, arising from the generalised pythagorean theorem.
- $[0,\infty] :=$ a category consisting of one object $\bullet$, with morphisms given by the elements of the set $\mathbb{R}^+ \cup \{\infty\}$, and their composition defined by addition (Lawvere'73).
- $2 :=$ category consisting of two objects, one arrow between them, and the identity arrows on each of the objects.
- $[0,\infty]^2$ has objects given by morphisms of $[0,\infty]$, morphisms given by the commutative squares in $[0,\infty]$, and compositions given by commutative compositions of these squares.
- Let $\mathtt{Aff}^{\subseteq}_{Q}(\ell, f)$ denote a full subcategory of $\mathtt{Aff}^{\subseteq}(\ell, f)$, determined by the choice of its terminal object to be given by $Q \in \mathrm{Ob}(\mathtt{Aff}^{\subseteq}(\ell, f))$.

Brègman relative entropy as a functor (II)

- Let $K_1, K_2, K_3, K, L \in \mathrm{Ob}(\mathtt{Aff}^{\subseteq}_{Q}(\ell, f))$, $K \subseteq K_2$ and $L \subseteq K_3$.
- For each $\phi \in Q$, the generalised pythagorean theorem implies the commutativity of the diagram

$$\begin{array}{ccc}
\bullet & \xrightarrow{\tilde{D}_f(\phi,x)} & \bullet \\
\uparrow{\scriptstyle 0} & & \uparrow{\scriptstyle \tilde{D}_f(\mathfrak{P}^{\tilde{D}_f}_K(x),x)} \\
\bullet & \xrightarrow{\tilde{D}_f(\phi,\mathfrak{P}^{\tilde{D}_f}_K(x))} & \bullet \\
\uparrow{\scriptstyle 0} & & \uparrow{\scriptstyle \tilde{D}_f(\mathfrak{P}^{\tilde{D}_f}_L \circ \mathfrak{P}^{\tilde{D}_f}_K(x),\mathfrak{P}^{\tilde{D}_f}_K(x))} \\
\bullet & \xrightarrow{\tilde{D}_f(\phi,\mathfrak{P}^{\tilde{D}_f}_L \circ \mathfrak{P}^{\tilde{D}_f}_K(x))} & \bullet
\end{array} \tag{3}$$

which implies the commutativity of

$$\begin{array}{ccccc}
x & \longmapsto & (\;\bullet & \xrightarrow{\tilde{D}_f(\phi,x)} & \bullet\;) \\
\downarrow{\scriptstyle \mathfrak{P}^{\tilde{D}_f}_K} & & \uparrow{\scriptstyle 0} & & \uparrow{\scriptstyle \tilde{D}_f(\mathfrak{P}^{\tilde{D}_f}_K(x),x)} \\
\mathfrak{P}^{\tilde{D}_f}_K(x) & \longmapsto & (\;\bullet & \xrightarrow{\tilde{D}_f(\phi,\mathfrak{P}^{\tilde{D}_f}_K(x))} & \bullet\;) \\
\downarrow{\scriptstyle \mathfrak{P}^{\tilde{D}_f}_L} & & \uparrow{\scriptstyle 0} & & \uparrow{\scriptstyle \tilde{D}_f(\mathfrak{P}^{\tilde{D}_f}_L \circ \mathfrak{P}^{\tilde{D}_f}_K(x),\mathfrak{P}^{\tilde{D}_f}_K(x))} \\
\mathfrak{P}^{\tilde{D}_f}_L \circ \mathfrak{P}^{\tilde{D}_f}_K(x) & \longmapsto & (\;\bullet & \xrightarrow{\tilde{D}_f(\phi,\mathfrak{P}^{\tilde{D}_f}_L \circ \mathfrak{P}^{\tilde{D}_f}_K(x))} & \bullet\;).
\end{array}$$

Brègman relative entropy as a functor (III)

- This defines a contravariant functor $\tilde{D}_f(\phi,\cdot):\mathtt{Aff}^{\subseteq}_{Q}(\ell,f)\to[0,\infty]^2$.
- It naturally extends to the functor $\tilde{D}_f(\phi,\cdot):\mathtt{Aff}^{\subseteq}(\ell,f)\downarrow Q\to[0,\infty]^2$, where $\mathtt{Aff}^{\subseteq}(\ell,f)\downarrow Q$ denotes a slice category of $\mathtt{Aff}^{\subseteq}(\ell,f)$ over Q.
- For any two categories $\mathtt{C}$ and $\mathtt{D}$, the cartesian closedness of the category $\mathtt{Cat}$ of all small categories (with natural transformations as morphisms) implies that any functor $\mathtt{C}\to\mathtt{D}^2$ corresponds to a natural transformation in the functor category $\mathtt{D}^{\mathtt{C}}$.
- Hence, Q parametrises the family of natural transformations $\tilde{D}_f(\phi,\cdot)$ in the category of functors $\mathtt{Aff}^{\subseteq}(\ell,f)\downarrow Q\to[0,\infty]$.

Resource theoretic view

- Given any object Q in $\mathtt{Cvx}(\ell, f)$, the set $\hom_{\mathtt{Cvx}(\ell,f)}(\,\cdot\,, Q)$ can be equipped with the structure of **commutative ordered monoid** via:
  - $\mathfrak{P}^{\tilde{D}_f}_{Q_1} \wedge \mathfrak{P}^{\tilde{D}_f}_{Q_2} := \mathfrak{P}^{\tilde{D}_f}_{Q_1 \cap Q_2}$,
  - $\mathfrak{P}^{\tilde{D}_f}_{Q_1} \leq \mathfrak{P}^{\tilde{D}_f}_{Q_2} := Q_1 \subseteq Q_2$,
  - distinguished zero object given by $\mathfrak{P}^{\tilde{D}_f}_{Q}$.
- Hence, each $\hom_{\mathtt{Cvx}(\ell,f)}(\,\cdot\,, Q)$ forms a resource theory in the sense of Fritz'17.
- Example: For $\tilde{D}_{1/2}$ defined by $X =$ Hilbert space $\mathcal{H}$, $\ell(\rho) = \rho^{1/2}$, $f = \frac{1}{2}\|\cdot\|^2_{\mathcal{H}}$ and under restriction to such Q that correspond to closed linear subspaces of $\mathcal{H}$, the projections $\mathfrak{P}^{\tilde{D}_f}_{Q}$ are given by the Hilbert space projection operators, while the operator implementing the finite join operation $\mathfrak{P}^{\tilde{D}_f}_{Q_1} \wedge \ldots \wedge \mathfrak{P}^{\tilde{D}_f}_{Q_n}$ is given by the von Neumann'49–Kakutani'40–Halperin'62 theorem:

$$\lim_{k\to\infty} \left\| ((P_{Q_n} \cdots P_{Q_1})^k - P_{Q_1 \cap \ldots \cap Q_n})\xi \right\|_{\mathcal{H}} = 0 \;\; \forall \xi \in \mathcal{H}.$$