Notes on algebraic information theory

Galobel Wang
The University of Hong Kong

May 2025

Abstract

This article consists of some notes taken by the author while studying algebraic information theory.

References: [1], [2]

Contents

1 Introduction

1.1 Entropies and their algebraic characterization.

Shannon [3] defined the information content of a random variable X, taking values in a finite set $\mathcal{E}_X$, by the formula

$$S_1[X](P) := -\sum_{x\in\mathcal{E}_X} P(X=x)\log P(X=x), \tag{1.1}$$

where P denotes a probability measure (law) on $\mathcal{E}_X$. The function S_1 is called **(Gibbs-Shannon) entropy**, and quantifies the uncertainty of a measurement.

Given two random variables X and Y, valued respectively in sets $\mathcal{E}_X$ and $\mathcal{E}_Y$, their joint measurement (X,Y) is also random variable, valued in $\mathcal{E}_{XY} \subset \mathcal{E}_X \times \mathcal{E}_Y$. The probability of observing $X = x$ is computed as the sum of all the outputs of (X,Y) that contain x in the first component: $X_*P(x) := P(X = x) = \sum_{(x,y)\in\mathcal{E}_{XY}} P(x,y)$. The probability X_*P on $\mathcal{E}_X$ is called **marginal law**. Instead of measuring directly (X,Y) one could measure first X, which constitutes a first random choice; the uncertainty that remains after obtaining the result $X = x_0$ is represented by the **conditional probability law**$P|_{X=x_0} : \mathcal{E}_{XY} \to [0,1]$, given by

$$P|_{X=x_0}(x,y) := \begin{cases} \frac{P(x,y)}{X_*P(x_0)} & \text{if } x = x_0 \\ 0 & \text{otherwise} \end{cases},$$

provided $X_*P(x_0) > 0$ (it remains undefined for x_0 in the maximal X_*P-null set). The function S_1 satisfies the so-called **chain rule**

$$S_1[(X,Y)](P) = S_1[X](X_*P) + \sum_{\substack{x\in\mathcal{E}_X \\ X_*P(x)>0}} X_*P(x) S_1[Y](X_*P|_{X=x}) \tag{1.2}$$

Evidently, if the measurement of Y is performed first, we obtain another equation, that corresponds to

$$S_1[(X,Y)](P) = S_1[Y](Y_*P) + \sum_{\substack{y\in\mathcal{E}_Y \\ Y_*P(y)>0}} Y_*P(y) S_1[X](X_*P|_{Y=y}) \tag{1.3}$$

Shannon [3] gave an algebraic characterization of

$$H_n : \Delta^n \to \mathbb{R}, \quad (p_0, ..., p_n) \mapsto -\sum_{i=0}^{n} p_i \log p_i$$

as the only family of continuous functions that satisfies the chain rule, for arbitrary pairs (X,Y), setting $S_1[X] = H_{|\mathcal{E}_X|}$ and so on—and such that $H_n(1/n, ..., 1/n)$ is monotonic in n.

In the same vein, if the product (X,Y) is nondegenerate (see below), then the system of functional equations (1.2)-(1.3), with measurable *unknowns* $S_1[X]$, $S_1[Y]$, and $S_1[(X,Y)]$, is uniquely solved by the corresponding Shannon entropies (1.1), up to a multiplicative constant.

More importantly, the chain-rule-like functional equations (1.2)-(1.3) accept a cohomological interpretation. Let us define, for any probabilistic functional $P \mapsto f(P)$, a new functional $X.f$ given by

$$(X.f)(P) := \sum_{x \in \mathcal{E}_X} X_*P(X) f(Y_*P|_{X=x}).$$

in order to rewrite (1.2) as

$$0 = X.S_1[Y] - S_1[(X,Y)] + S_1[X]. \tag{1.4}$$

The notation is meant to suggest an action of random variables on probabilistic functionals, and in fact the equality $Z.(X.f) = (Z,X).f$ holds. There is an strong resemblance between (1.4) and a cocycle equation in group cohomology. Baudot and Bennequin [2] formalized this analogy introducing an adapted cohomology theory—information cohomology—through an explicit differential complex that recovered the equations (1.4) as 1-cocycle conditions. They used presheaves, exploiting a notion of **locality** specific to the problem: the entropy of a variable X only depends on the marginalized version X_*P of any global law P.

1.2 Categories of observables

Information cohomology was introduced in [2] considering presheaves on **information structures**, that were either categories of partitions of a given measurable space or categories of orthogonal decompositions of a Hilbert space. The partitions corresponded to atoms of the σ-algebras generated by measurable functions (classical observables) with finite range, and the orthogonal decompositions appeared as eigenspaces of self-adjoint operators (quantum observables) with finitely many different eigenvalues.

We wanted to approach measurements from a categorical viewpoint, describing directly the relations between their outputs and without presupposing the existence of an underlying probability space or Hilbert space. A probability space is only necessary to represent a collection of observables by measurable functions with a common domain, as customary in "classical" probability theory (as opposed to "quantum"). The sets of outputs can also be interpreted as the spectra of self-adjoint operators, in such a way that some contextual collections have *quantum representations*.

In view of the foregoing, we introduce here a more general definition of information structure, that covers the classical and quantum cases at the same time and extends without modification to continuous random variables. This allows us to introduce a category of information structures and to treat the algebraic aspects of the theory in a unified manner, once for all these cases. To attain this flexibility and generality, the definition decouples the combinatorial structure of joint measurements and the local models of the outputs of each individual measurement.

Let $\mathbf{Meas}_{\text{surj}}$ be the category of measurable spaces and measurable surjections between them.

Definition 1.1 (conditional meet semilattice is a poset)**.** *A conditional meet semilattice is a poset that satisfies the following property:*

for any $X, Y, Z \in Ob\,\mathbf{S}$, if $Z \to X$ and $Z \to Y$, then the categorical product $X \wedge Y$ exists.

It is unital whenever it has a terminal object, denoted $\top$.

Definition 1.2 (information structure)**.** *An information structure is a pair $(\mathbf{S}, \mathcal{M})$, where $\mathbf{S}$ is a **unital conditional meet semilattice** and $\mathcal{M} : \mathbf{S} \to \mathrm{Meas}_{\mathrm{surj}}$ is a functor (say $\mathcal{M}(X) = \mathcal{M}_X = (\mathscr{E}_X, \mathscr{B}_X)$, for each $X \in \mathrm{Ob}\,\mathbf{S}$) that satisfies:*

1. *$\mathscr{E}_\top \cong \{*\}$, with the trivial σ-algebra;*
2. *for every $X \in \mathrm{Ob}\,\mathbf{S}$ and any $x \in \mathscr{E}_X$, the σ-algebra $\mathscr{B}_X$ contains the singleton $\{x\}$;*
3. *for every diagram $X \xleftarrow{\pi} X \wedge Y \xrightarrow{\sigma} Y$ the measurable map*

$$\mathcal{M}_{X\wedge Y} \hookrightarrow \mathcal{M}_X \times \mathcal{M}_Y,\; z \mapsto (x(z), y(z)) := (\mathcal{M}_\pi(z), \mathcal{M}_\sigma(z))$$

is an injection.

The objects of the **conditional meet semilattice S** stand for observables and the arrows encode the relation of **refinement** between them (think of refinements of σ-algebras or orthogonal decompositions). The terminal object is "certainty", the meet $X \wedge Y$ represents the joint measurement of X and Y, and condition of categorical product accommodates the impossibility of doing some joint measurements. For instance, in quantum mechanics, it is only possible to jointly measure X and Y if they commute, in which case the observable (X, Y) induces an orthogonal decomposition of the Hilbert space that refines the decompositions induced by X and Y. In turn, the functor $\mathcal{M}$ represents the possible outputs of each observable. A refinement $\pi : X \to Y$ translates into a surjection $\pi_* \equiv \mathcal{M}\pi : \mathcal{M}_X \to \mathcal{M}_Y$ that induces an injection at the level of the algebras of events $\pi^* : \mathcal{B}_Y \to \mathcal{B}_X$ that maps A to $\mathcal{M}\pi^{-1}(A)$, compare with the extensions of probability spaces. The set $\mathcal{E}_{X\wedge Y}$ represents the possible outputs of the joint measurement $X \wedge Y$, hence it can be identified with a subset of $\mathcal{E}_X \times \mathcal{E}_Y$. When convenient, we use the notations common in probability theory: $\{X = x\}$ means "the element x contained in $\mathcal{E}_X$" and $\{X = x, Y = y\}$ should be interpreted as the element z of $\mathcal{E}_{X\wedge Y}$ mapped to x by $\mathcal{E}_{X\wedge Y} \to \mathcal{E}_X$ and to y by $\mathcal{E}_{X\wedge Y} \to \mathcal{E}_Y$ (if such z does not exist, $\{X = x, Y = y\} = \varnothing$).

Probability laws come as a functor $\mathscr{P} : \mathbf{S} \to \mathbf{Sets}$ that associates to each $X \in \mathrm{Ob}\,\mathbf{S}$ the set $\mathscr{P}_X$ of measures P on $\mathcal{M}_X$ such that $P(\mathcal{E}_X) = 1$. Each arrow $\pi : X \to Y$ induces a measurable surjection $\mathcal{M} : \mathcal{M}_X \to \mathcal{M}_Y$, and $\mathscr{P}\pi : \mathscr{P}_X \to \mathscr{P}_Y$ is defined to be the push-forward of measures: for every $B \in \mathcal{B}_Y$,

$$(\mathscr{P}\pi(P))(B) = P(\mathcal{M}\pi^{-1}(B)).$$

This operation is called **marginalization**. We write π_* or Y_* instead of $\mathscr{P}\pi$, if there is no risk of ambiguity; this notation is compatible with that of previous subsection.

1.3 An example: simplicial information structures

Example 1.1. *If I is any set, let $\Delta(I)$ be the poset of its finite subsets, with an arrow $A \to B$ whenever $B \subset A$. A simplicial subcomplex of $\Delta(I)$ is a full subcategory $\mathbf{K}$ such that, for any given object of $\mathbf{K}$ ("a cell"), all its subsets are also objects of $\mathbf{K}$ ("faces"). Given a collection $\{(\mathcal{E}_i, \mathcal{B}_i)\}_{i\in I}$ of masurable spaces, let $\mathcal{M} : \Delta(I) \to \mathbf{Meas_surj}$ be the functor that associates to each $A \subset I$ the set $\mathcal{E}_A := \prod_{i\in A} \mathcal{E}_i$ with the product σ-algebra $\mathcal{B}_A := \bigotimes_{i\in A} \mathcal{B}_i$, and to each arrow in $\Delta(I)$ the corresponding canonical projector. The pair $(\mathbf{K}, \mathcal{M}|_{\mathbf{K}})$ is a **simplicial information structure**.*

2 The category of information structures

2.1 Terminology and examples

Here we call random variables (r.v) on a finite set Ω **congruent** when they define the same partition (remind that a partition of Ω is a family of disjoint non-empty subsets covering Ω and that the partition associated to a r.v X is the family of subsets $\Omega_x \subset \Omega$ defined by the equations $X(\omega) = x$).

Let Ω be a finite set, the set $\Pi(\Omega)$ of all partitions of Ω constitutes a category with one arrow $Y \to Z$ from Y to Z when Y is more fine than Z, we also say in this case that Y divides Z.

In $\Pi(\Omega)$ we have an **initial element**, which is the partition by points, denoted ω and a **final(terminal) element**, which is Ω itself and is denoted by 1.

The **joint partition** YZ or (Y, Z), of two partitions Y, Z of Ω is the less fine partition that divides Y and Z, i.e., their gcd.

For any X we get $XX = X$, $\omega X = \omega$ and $1.X = X$, which implies each partition is idempotent.

From above properties of the partitions, we can strictly give the definition of **information structure**.

Definition 2.1 (Partially ordered set)**.** *A **reflexive, weak** or **non-strict partial order** commonly referred to simply as a **partial order**, is a homogeneous relation $\leq$ on a set P that is reflexive, antisymmetric, and transitive. That is, for all $a, b, c \in P$, it must satisfy:*

1. ***Reflexivity****: $a \leq a$, i.e. every element is related to itself.*

2. ***Antisymmetry****: if $a \leq b$ and $b \leq a$ then $a = b$, i.e. no two distinct elements precede each other.*

3. ***Transitivity****: if $a \leq b$ and $b \leq c$ then $a \leq c$.*

Definition 2.2 (meet-semilattice)**.** *A set S partially ordered by the binary relation $\leq$ is a* ***meet-semilattice*** *if for all elements x and y of S, the greatest lower bound of the set $\{x, y\}$ exists.*

The greatest lower bound of the set $\{x, y\}$ is called the ***meet*** *of x and y, denoted $x \wedge y$.*

Replacing "greatest lower bound" with "least upper bound" results in the dual concept of a **join-semilattice**. The least upper bound of $\{x, y\}$ is called the **join** of x and y, denoted $x \vee y$.

By definition an **information structure** $\mathcal{S}$ on Ω is a subset of $\Pi(\Omega)$, such that for any element X of $\mathcal{S}$, and any pair of elements Y, Z in $\mathcal{S}$ that X refines, the joint partition YZ also belongs to $\mathcal{S}$. In addition we will always assume that the final partition 1 belongs to $\mathcal{S}$. In terms of observations, it means that at least something is a certitude.

Definition 2.3 (conditional meet semilattice is a poset)**.** *A conditional meet semilattice is a poset that satisfies the following property:*

for any $X, Y, Z \in Ob\,\mathbf{S}$, if $Z \to X$ and $Z \to Y$, then the categorical product $X \wedge Y$ exists.

It is unital whenever it has a terminal object, denoted $\top$.

Definition 2.4 (information structure)**.** *An information structure is a pair $(\mathbf{S}, \mathcal{M})$, where $\mathbf{S}$ is a* ***unital conditional meet semilattice*** *and $\mathcal{M} : \mathbf{S} \to \mathrm{Meas}_{\mathrm{surj}}$ is a functor (say $\mathcal{M}(X) = \mathcal{M}_X = (\mathscr{E}_X, \mathscr{B}_X)$, for each $X \in \mathrm{Ob}\,\mathbf{S}$) that satisfies:*

1. *$\mathscr{E}_\top \cong \{*\}$, with the trivial σ-algebra;*

2. *for every $X \in \mathrm{Ob}\,\mathbf{S}$ and any $x \in \mathscr{E}_X$, the σ-algebra $\mathscr{B}_X$ contains the singleton $\{x\}$;*

3. *for every diagram $X \xleftarrow{\pi} X \wedge Y \xrightarrow{\sigma} Y$ the measurable map*

$$\mathcal{M}_{X \wedge Y} \hookrightarrow \mathcal{M}_X \times \mathcal{M}_Y,\; z \mapsto (x(z), y(z)) := (\mathcal{M}_\pi(z), \mathcal{M}_\sigma(z))$$

   is an injection.

An information structure $(\mathbf{S}, \mathcal{M})$ is said to be **bounded** if the poset $\mathbf{S}$ has finite height. It is **finite** if all the sets $\mathscr{E}_X$ are finite, in which case $\mathscr{E}_X$ corresponds to the atoms of $\mathscr{B}_X$ and the algebra can be omitted from the description. We denote a finite structure by $(\mathbf{S}, \mathscr{E})$, where $\mathscr{E}$ is a covariant functor from $\mathbf{S}$ to $\mathbf{Sets}$. The cohomological computations concern finite structures, but the general constructions do not require this hypothesis. In fact, they only depend on the combinatorial object $\mathbf{S}$.

Example 2.1. *Start with a set $\Sigma = \{S_i; 1 \leq i \leq n\}$ of partitions of Ω. For any subset $I = \{i_1, ..., i_k\}$ of $[n] = \{1, ..., n\}$, the joint $(S_{i_1}, ..., S_{i_k})$, also denoted S_I, divides each S_{i_j}. The set $W = W(\Sigma)$ of all the S_I, when I describes the subsets of $[n]$ is an information struture. It is even a commutative monoid, because any product of elements of W belongs to W, and the partition associated to Ω itself gives the identity element of W. The product $S_{[n]}$ of all the S_i is maximal; it divides all the other elements. As $\Pi(\Omega)$ the monoid $W(\Sigma)$ is idempotent, i.e., for any X we have $XX = X$.*

Example 2.2 (Concrete structures)**.** *Given a set Ω, let $\mathbf{Obs}_{\text{fin}}(\Omega)$ be the category **finite observables**; the objects of this category are finite partitions of Ω, and there is an arrow $X \to Y$ whenever X refines Y. In this case, X discriminates better between the **configurations** $w \in \Omega$. The category $\mathbf{Obs}_{\text{fin}}(\Omega)$ has a terminal object: the trivial partition $\top := \{\Omega\}$. When Ω is finite, it also has an initial object: the partition by points, that we denote by $\perp$. The categorical product $X \times Y$ of two partitions X and Y is the coarsest partition that refines both. This product is commutative, associative, idempotent and unitary ($\top \times X = X$).*

A **classical information structure** in the sense of [2] is a **full subcategory $\mathbf{S}$** of $\mathbf{Obs}_{\text{fin}}(\Omega)$ such that

- $\top \in \operatorname{Ob} \mathbf{S}$;
- for any X, Y, Z in $\operatorname{Ob} \mathbf{S}$, if $X \to Y$ and $X \to Z$, then $Y \times Z$ belongs to $\mathbf{S}$.

We call $\mathbf{S}$ a **concrete structure**. If $\square : \mathbf{Obs}_{\text{fin}}(\Omega) \to \mathbf{Sets}$ denotes the "forgetful" functor that maps the partition $X = \{A_1, ..., A_n\}$ to the set $\{A_1, ..., A_n\}$ and each arrow $X \to Y$ in $\mathbf{Obs}_{\text{fin}}(\Omega)$ to the unique surjective map $\square\pi : \mathscr{E}_X \to \mathscr{E}_Y$ such that $B = \bigcup_{A \in \mathscr{E}_{\pi^{-1}(B)}} A$ for any $B \in \mathscr{E}_Y$, the pair $(\mathbf{S}, \square)$ is a finite information structure according to the definition of information structure.

Concrete structures turn out to be too restrictive. For instance, Baudot and Bennequin [2] associate to any finite indexed collection $\Sigma = (S_1, ..., S_n)$ of partitions of Ω a **simplicial structure $\mathbf{S}(\mathbf{K})$**: a subcategory of $\mathbf{Obs}_{\text{fin}}(\Omega)$ that contains $\prod_{i \in A} S_i$ for any object A of a simplicial subcomplex $\mathbf{K}$ of the abstract simplex $\Delta(\{1, ..., n\})$; by convention, the empty product gives the trivial partition. Such construction does not necessarily give an information structure (in their sense). For example: if $n = 3$, $\Omega = \{0, 1\}^2$, S_i is the partition induced by the projection on the i-th component $(i = 1, 2)$, $S_3 = \{\{(0, 0)\}, \{(0, 1)\}\}$, and the maximal cells of $\mathbf{K}$ are $\{1, 2\}$ and $\{3\}$, then $S_1 \times S_2$ is the atomic partition, that refines all the others, while some products (like $S_1 \times S_3$) are not in $\mathbf{S}(\mathbf{K})$. In our framework, the category $\mathbf{S}(\mathbf{K})$ appears as a classical representation of the (generalized) information structure $(\mathbf{K}, \mathscr{E})$, where $\mathscr{E} : \mathbf{K} \to \mathbf{Sets}$ is given by $\mathscr{E}_{\{1\}} = \mathscr{E}_{\{2\}} = \{0, 1\}$, $\mathscr{E}_{\{1,2\}} = \mathscr{E}_{\{1\}} \times \mathscr{E}_{\{2\}}$, the maps induced by the arrows in $\mathbf{K}$ being canonical projections.

Example 2.3 (Homogeneous structures). *Let G be a locally compact, Hausdorff topological group. Any collection $\mathcal{C}$ of closed subgroups of G that contains G and is conditionally closed under intersections (i.e. for any $M, N, O \in \mathcal{C}$, if $N \subset M$ and $N \subset O$, then $M \cap O \in \mathcal{C}$) defines a conditional meet semilattice $\mathbf{S}$, whose arrows correspond to inclusions. Let $\mathcal{M}$ be the functor that associates to each subgroup N the (Hausdorff) quotient space G/N with the Borel σ-algebra induced by the quotient topology, and to each arrow $N \to M$ the canonical projection $\pi_{M,N} : G/N \to G/M$ that sends the coset gN to gM. The information structures $(\mathbf{S}, \mathcal{M})$ obtained in this way are called **homogeneous**, because each coset space G/M is a homogeneous space for G.*

2.2 Idempotent monoids and sheaf theory

Definition 2.5. *A monoid $(M, \cdot, e)$ is **idempotent** if for all $m \in M$, $m \cdot m = m$.*

Any conditional meet semilattice $\mathbf{S}$ induces a **presheaf** of idempotent monoids on it: for each $X \in \operatorname{Ob} \mathbf{S}$, set $\mathscr{I}_X := \{Y \in \operatorname{Ob} \mathbf{S} \mid X \to Y\}$, with the monoid structure given by the product of observables in $\mathbf{S}$: $(Z, Y) \mapsto ZY := Z \wedge Y$; an arrow $X \to Y$ in $\mathbf{S}$ induces an inclusion $\mathscr{I}_Y \hookrightarrow \mathscr{I}_X$.

Definition 2.6 (presheaf). *Let X be a topological space. A **presheaf of groups** $\mathcal{F}$ on X is a function which assigns to every open set $U \subset X$ a group $\mathcal{F}(U)$ and to every inclusion $V \subset U$ a restriction map,*

$$\rho_{UV} : \mathcal{F}(U) \longrightarrow \mathcal{F}(V),$$

which is a group homomorphism, such that if $W \subset V \subset U$, then

$$\rho_{VW} \circ \rho_{UV} = \rho_{UW}.$$

Succinctly put, a pre-sheaf is a contravariant functor from $\mathbf{Top}(X)$ to the category (**Groups**) of groups. Put this way, it is clear what we mean by a presheaf of rings, etc. The elements of $\mathcal{F}(U)$ are called **sections**. We almost always denote $\rho_{UV}(s) = s|_V$. U_{ij} denotes $U_i \cap U_j$.

Example 2.4. *Let X be a topological space and let G be a group. Define a presheaf $\mathcal{G}$ as follows. Let U be any open subset of X. $\mathcal{G}(U)$ is defined to be the set of constant functions from X to G. The restriction maps are the obvious ones.*

Definition 2.7 (sheaf). *A **sheaf** $\mathcal{F}$ on a **topological space** is a presheaf which satisfies the following two axioms.*

1. *Given an open cover U_i of U an open subset of X, and a collection of sections s_i on U_i, such that $s_i|_{U_{ij}} = s_j|_{U_{ij}}$ then there is a section s on U such that $s|_{U_i} = s_i$.*
2. *Given an open cover U_i of U an open subset of X, if s is a section on U such that $s|_{U_i} = 0$, then s is zero.*

Note that we could easily combine axioms (1) and (2) and require that there is a unique s, which is patched together from the s_i. It is very easy to give lots of examples of sheaves and presheaves. Basically, any collection of functions is a sheaf.

Example 2.5. *Let M be a complex manifold. Then there are a collection of sheaves on M. The sheaf of holomorphic functions, the sheaf of C^∞-functions and the sheaf of continuous functions. In all cases, the restrictions maps are the obvious ones, and there are obvious inclusions of sheaves.*

Given a variety X, the sheaf of regular functions is a sheaf of rings.

Note however that in general the presheaf defined in example 2.4 is not a sheaf. For example, take $X = \{a, b\}$ to be the topological space with the discrete topology and take $G = \mathbb{Z}$. Let $U_1 = \{a\}$ and $U_2 = \{b\}$ and suppose $s_1 = 0$ and $s_2 = 1$. Then there is no global constant function which restricts to both 0 and 1.

However this is easily fixed. Take $\mathcal{F}$ to be the sheaf of locally constant functions.

Definition 2.8 (stalk)**.** *Let X be a topological space and let $\mathcal{F}$ be a presheaf on X. Let $p \in X$. The **stalk** of $\mathcal{F}$ at p, denoted $\mathcal{F}_p$, is the inverse limit*

$$\varinjlim_{p \in U} \mathcal{F}(U).$$

It is useful to untwist this definition. An element of the stalk is a pair (s, U), such that $s \in \mathcal{F}(U)$, modulo the equivalence relation,

$$(s, U) \sim (t, V)$$

if there is an open subset $W \subset U \cap V$ such that

$$s|_W = t|_V.$$

In other words, we only care about what s looks like in an arbitrarily small neighbourhood of p. Note that when we have a sheaf of rings, the stalk is often a local ring.

Definition 2.9. *A map between presheaves is a natural transformation of the corresponding functors.*

Untwisting the definition, a map between presheaves

$$f : \mathcal{F} \longrightarrow \mathcal{G}$$

assigns to every open set U a group homomorphism

$$f(U) : \mathcal{F}(U) \longrightarrow \mathcal{G}(U),$$

such that the following diagram always commutes

$$\begin{tikzcd}
\mathcal{F}(U) \arrow[r, "f(U)"] \arrow[d, "\rho_{UV}"'] & \mathcal{G}(U) \arrow[d, "\sigma_{UV}"] \\
\mathcal{F}(V) \arrow[r, "f(V)"'] & \mathcal{G}(V)
\end{tikzcd}$$

Note that this gives us a category of presheaves, together with a full subcategory of sheaves.

Lemma 2.1. *Let $\mathcal{F}$ be a presheaf. Then the* ***sheaf associated to the presheaf****, is a sheaf $\mathcal{F}^+$, together with a morphism of sheaves $u : \mathcal{F} \longrightarrow \mathcal{F}^+$ which is universal amongst all such morphisms of sheaves: that is given any morphism of presheaves*

$$f : \mathcal{F} \longrightarrow \mathcal{G},$$

where $\mathcal{G}$ is a sheaf, there is a unique induced morphism of sheaves which makes the following diagram commute

$$\begin{tikzcd}
\mathcal{F} \arrow[r, "f"] \arrow[d, "u"'] & \mathcal{G} \\
\mathcal{F}^+ \arrow[ur, dashed] &
\end{tikzcd}$$

Proposition 2.1. *Let $\phi : \mathcal{F} \longrightarrow \mathcal{G}$ be a morphism of sheaves. Then ϕ is an isomorphism iff the induced map on stalks is always an isomorphism.*

Proof. One direction is clear. So suppose that the map on stalks is an isomorphism. It suffices to prove that $\phi(U) : \mathcal{F}(U) \longrightarrow \mathcal{G}(U)$ is an isomorphism, for every open subset $U \subset X$, since then the inverse morphism ϕ is given by setting $\psi(U) = \phi(U)^{-1}$.

We first show that $\phi(U)$ is injective. Let $s \in \mathcal{F}(U)$ and suppose that $\phi(U)(s) = 0$. Then surely $\phi_p(s_p) = 0$, where $s_p = (s, U) \in \mathcal{F}_p$ and $p \in U$ is arbitrary. Since ϕ_p is injective by assumption, it follows that there is an open set $V_p \subset U$ containing p such that $s|_{V_p} = 0$. $\{V_p\}_{p \in U}$ is an open cover of U and as $\mathcal{F}$ is a sheaf, it follows that $s = 0$. Hence $\phi(U)$ is injective, for every U.

Now we show that $\phi(U)$ is surjective. Suppose that $t \in \mathcal{F}(U)$. Since ϕ_p is surjective, for every p, we may find an open set $p \in U_p \subset U$ and a section $s_p \in \mathcal{F}(U_p)$ such that $\phi(U_p)(s_p) = t|_{U_p}$. Pick p and $q \in U$ and set $V = U_p \cap U_q$. Then $\phi(V)(s_p|_V) = \phi(V)(s_q|_V)$. Since $\phi(V)$ is injective, it follows that $s_p|_V = s_q|_V$. As $\mathcal{F}$ is a sheaf, it follows that there is a section $s \in \mathcal{F}(U)$ such that $\phi(U)(s) = t$. But then $\phi(U)$ is surjective. $\square$

Example 2.6. *Let $X = \mathbb{C} - \{0\}$, let $\mathcal{F} = \mathcal{O}_X$, the sheaf of holomorphic functions and let $\mathcal{G} = \mathcal{O}_X^*$, the sheaf of non-zero holomorphic functions.*

There is a natural map

$$\phi : \mathcal{F} \longrightarrow \mathcal{G},$$

which just sends a function f to its exponential. Then ϕ is surjective on stalks; this just says that given a non-zero holomorphic function g, then $\log(g)$ makes sense in a small neighbourhood of any point.

Definition 2.10 (push-forward)**.** *Let $f : X \longrightarrow Y$ be a continuous map of topological spaces. Let $\mathcal{F}$ be a sheaf on X. The **pushforward of** $\mathcal{F}$, denoted $f_*\mathcal{F}$, is defined as follows*

$$f_*\mathcal{F}(U) = \mathcal{F}(f^{-1}(U)),$$

where $U \subset Y$ is an open set.

Definition 2.11 (inverse image)**.** *Let $\mathcal{G}$ be a sheaf on Y. The **inverse image of** $\mathcal{G}$, denoted $f^{-1}\mathcal{G}$, is the sheaf assigned to the presheaf*

$$U \longmapsto \varinjlim_{f(U) \subset V} \mathcal{G}(V),$$

where U is an open subset of X and V ranges over all open subsets of Y which contain $f(U)$.

Definition 2.12 (ringed space)**.** *A pair $(X, \mathcal{O}_X)$ is called a **ringed space**, if X is a topological space, and $\mathcal{O}_X$ is a sheaf of (commutative) rings. A **morphism** $\phi : X \longrightarrow Y$ of ringed spaces is a pair $(f, f^\#)$, consisting of a continuous function $f : X \longrightarrow Y$ and a sheaf morphism $f^\# : \mathcal{O}_Y \longrightarrow f_*\mathcal{O}_X$.*

Definition 2.13. *A **locally ringed space**, is a ringed space $(X, \mathcal{O}_X)$ such that in addition every stalk $\mathcal{O}_{X,x}$ of the structure sheaf is a local ring. A morphism of locally ringed spaces is a morphism of ringed spaces, such that for every point $x \in X$, the induced map*

$$f_x^\# : \mathcal{O}_{Y,y} \longrightarrow \mathcal{O}_{X,x},$$

is a morphism of local rings (that is the inverse image of the maximal ideal of $\mathcal{O}_{X,x}$ is a subset of the maximal ideal of $\mathcal{O}_{Y,y}$, where $y = f(x)$).

Note that we get a category of ringed spaces, whose objects are ringed spaces and whose morphisms are morphisms of ringed spaces. Further the category of locally ringed spaces is a subcategory.

Definition 2.14. *Let $(X, \mathcal{O}_X)$ be a ringed space. An $\mathcal{O}_X$-**module** is a sheaf $\mathcal{F}$ such that for every open set $U \subset X$, $\mathcal{F}(U)$ has the structure of an $\mathcal{O}_X(U)$-module, compatible with the restriction map, in an obvious way.*

By lemma 2.1 we may define various natural operations on sheaves. For example, let $\mathcal{F}$ and $\mathcal{G}$ be two $\mathcal{O}_X$-modules. The tensor product of $\mathcal{F}$ and $\mathcal{G}$, denoted $\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{G}$, is the sheaf associated to the presheaf

$$U \longmapsto \mathcal{F}(U) \otimes_{\mathcal{O}_X(U)} \mathcal{G}(U),$$

and curly hom, denoted $\mathbf{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$, is the sheaf associated to the presheaf

$$U \longmapsto \mathbf{Hom}_{\mathcal{O}_X(U)}(\mathcal{F}(U), \mathcal{G}(U)).$$

Let $f \colon \mathcal{F} \longrightarrow \mathcal{G}$ be a morphism of sheaves. The kernel of f is the sheaf which assigns to every open set U the kernel of the homomorphism $f(U) \colon \mathcal{F}(U) \longrightarrow \mathcal{G}(U)$. Similarly the image is the sheaf associated to the presheaf which assigns to every open set U the image of the homomorphism $f(U) : \mathcal{F}(U) \longrightarrow \mathcal{G}(U)$. We say that ϕ is injective iff $\mathrm{Ker}(\phi) = 0$ and we say that ϕ is surjective iff $\mathrm{Im}(\phi) = \mathcal{G}$.

Given a morphism of ringed spaces, and a sheaf $\mathcal{G}$ of $\mathcal{O}_X$-modules, the pullback of $\mathcal{G}$, denoted $\phi^*\mathcal{G}$, is the sheaf of $\mathcal{O}_Y$-modules,

$$\phi^{-1}\mathcal{G} \otimes_{f^{-1}\mathcal{O}_X} \mathcal{O}_Y.$$

Furthermore, there is a well-known equivalence between idempotent monoids and meet semilattices with a terminal object.

Proposition 2.2. *If $(M, \cdot, e)$ is an idempotent monoid, then the condition*

$$x \leq y \Leftrightarrow x \cdot y = x \tag{2.1}$$

defines a partial order on M such that any two elements of M have a meet and e is the greatest element.

Conversely, if $(E, \leq)$ is a poset with a greatest element in which any two elements $x, y \in E$ have a meet $x \wedge y$, then E endowed with the addition $(x, y) \mapsto x \wedge y$ is an idempotent monoid.

The two functors just introduced are inverses of each other.

Is there a counterpart to conditional meet semilattices with a terminal object in the theory of idempotent monoids? The following result serves as a partial answer. It involves **upper sets** of an idempotent monoid: a subset H of an idempotent monoid M—equipped with the partial order in (2.1)—is called an **upper set** if $h \in H$ and $h \leq m$ implies that $m \in H$. For example, the simplicial subcomplex $\mathbf{K}$ in example 1.1 defines an upper set of $\Delta(\mathbf{I})$, seen as an idempotent monoid according to Proposition 2.2.

Proposition 2.3. *Let $(M, \cdot, e)$ be an idempotent monoid, and $\mathbf{M}$ its associated poset* *(seen as a category). The full subcategory of* $\mathbf{M}$ *defined by any nonempty upper set* H *of* M *is a unital conditional meet semilattice.*

Proof. First, $e \in H$, because e is greater than any element of H. Second, if x, y, z be elements of H such that $z \leq x$ and $z \leq y$, then $z \leq x \wedge y$ in virtue of the universal property of $\wedge$ in $\mathbf{M}$, which in turn implies that $x \wedge y \in H$ (by definition of upper set). $\square$

2.3 Morphisms and (co)products

Definition 2.15 (pullback)**.** *Given category* $\mathcal{C}$ *with two morphisms* $f : B \to A$ *and* $g : C \to A$, *then a pullback is a triple* (D, α, β), *where* $\alpha : D \to C, \beta : D \to B$, *satisfying:*

1. $g\alpha = f\beta$
2. *Universal property: for all* (X, α', β'), $\alpha' : X \to C, \beta' : X \to B$ *and* $g\alpha' = f\beta'$, *exists unique morphism* $\theta : X \to D$ *such that* $\alpha' = \alpha\theta, \beta' = \beta\theta$, *and the following diagram commutes.*

$$\begin{tikzcd}
X \arrow[dr, dashed, "\theta"] \arrow[drr, bend left, "\alpha'"] \arrow[ddr, bend right, "\beta'"'] & & \\
& D \arrow[r, "\alpha"] \arrow[d, "\beta"] & C \arrow[d, "g"] \\
& B \arrow[r, "f"'] & A
\end{tikzcd}$$

Given categories $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$, and a functor $\xi : \mathbf{A} \to \mathbf{B}$, define the pullback $\xi^* : [\mathbf{B}, \mathbf{C}] \to [\mathbf{A}, \mathbf{C}]$ by $\mathscr{F} \mapsto \mathscr{F} \circ \phi$. It commutes with limits and colimits.

Definition 2.16. *A morphism* $\xi : \mathsf{S}_1 \to \mathsf{S}_2$ *of conditional meet semilattices is a functor (i.e. a monotone map) with the following property: if* $X \wedge Y$ *exists in* S_1, *then* $\xi(X \wedge Y) = \xi(X) \wedge \xi(Y)$.

A morphism $\phi : (\mathsf{S}, \mathscr{M}) \to (\mathsf{S}', \mathscr{M}')$ between information structures is a pair $\phi = (\phi_0, \hat{\phi})$ such that ϕ_0 is a morphism of conditional meet semilattices that maps T_S to $\mathsf{T}_{\mathsf{S}'}$, and $\phi : \mathscr{M} \to \phi_0^* \mathscr{M}'$ is a natural transformation. If there is no risk of ambiguity, we write ϕ instead of ϕ_0.

Given $\phi : (\mathsf{S}, \mathscr{M}) \to (\mathsf{S}', \mathscr{M}')$ and $\psi : (\mathsf{S}', \mathscr{M}') \to (\mathsf{S}'', \mathscr{M}'')$, their composition $\psi \circ \phi$ is defined as $(\psi_0 \circ \phi_0, \hat{\psi} \circ \hat{\phi} : \mathscr{M} \Rightarrow \psi_0^* \phi_0^* \mathscr{M}'')$.

We denote by **InfoStr** the category of information structures obtained in this way.

Note that, if $X \wedge Y$ exists, then $\xi(X \wedge Y) \to \xi(X)$ and $\xi(X \wedge Y) \to \xi(Y)$, and thus the product $\xi(X) \wedge \xi(Y)$ exists too.

A morphism of information structures is a particular case of morphism of $\mathbf{Meas}_{\text{surj}}$-valued covariant diagrams. We want ϕ_0 to respect the unit and the products, so that it induces a morphism between the corresponding presheaves of idempotent monoids.

The preceding definition is one of the main motivations for our generalized setting. In fact, one could imagine a correspondence between the partitions of two concrete structures defined on different sample spaces, but in which category would that correspondence take place? Since we eliminated the explicit reference to the sample space in our definition of information structure, the introduction of morphisms becomes straightforward. This allows the computation of products and coproducts.

Remark 2.1. *The connection to the sample spaces is not completely lost, but reformulated in the language of representations: if S_i is a concrete structure on Ω_i $(i = 1, 2)$, then the objects of $\mathsf{S}_1 \times \mathsf{S}_2$ can be identified with partitions of $\Omega_1 \times \Omega_2$, as one would expect.*

Proposition 2.4. *The category **InfoStr** has countable products and arbitrary coprod ucts.*

Proof. Let 0 be the category that has $\top$ as the only object and $\mathrm{id}_\top$ as the only morphism, and let $\mathcal{M}_0$ be the functor that associates to $\top$ the set $\{*\}$ equipped with the atomic σ-algebra. Clearly $(0, \mathcal{M}_0)$ is initial and terminal in the category **InfoStr**, hence it corresponds to the empty product and coproduct respectively.

Nonempty products: Given information structures $(\mathrm{S}_i, \mathcal{M}_i)$ indexed by i in an arbitrary set I, we introduce first the ordinary categorical product $\mathrm{S} = \prod_{i\in I} \mathrm{S}_i$: its objects are I-tuples $\langle X_i\rangle_{i\in I}$ with $X_i \in \mathrm{Ob}\,\mathrm{S}_i$ for each $i \in I$; there is an arrow $\langle \pi_i\rangle_{i\in I}$: $\langle X_i\rangle_{i\in I} \to \langle Y_i\rangle_{i\in I}$ whenever $\pi_i : X_i \to Y_i$ in S_i for each $i \in I$. Then a functor $\mathcal{M}$: $\mathrm{S} \to \mathrm{Meas}_{\mathrm{surj}}$ is defined as follows: for each $X = \langle X_i\rangle_{i\in I} \in \mathrm{Ob}\,\mathrm{S}$ the measurable space $\mathcal{M}(X)$ is the set $\mathcal{E}(X) := \prod_{i\in I} \mathcal{E}_i(X_i)$ equipped with the product σ-algebra $\mathcal{B}(X) := \bigotimes_{i\in I} \mathcal{B}_i(X_i)$, which is the smallest σ-algebra that makes every canonical projection $\widehat{p}^i_{\langle X_i\rangle_{i\in I}} : \mathcal{E}(\langle X_i\rangle_{i\in I}) \to \mathcal{E}_i(X_i)$ measurable; at the level of morphisms, $\mathcal{M}(\langle\pi_i\rangle_{i\in I}) := \prod_{i\in I} \mathcal{M}_i(\pi_i)$, which comes from the product in **Sets**.

The pair $(\mathrm{S}, \mathcal{M})$ is an information structure. It is easy to verify that S is a poset with terminal object $\langle \top_{S_i}\rangle_{i\in I}$. The conditional existence of products also holds: if $\langle X_i\rangle_{i\in I}, \langle Y_i\rangle_{i\in I}$ and $\langle Z_i\rangle_{i\in I}$ are objects of S such that $\langle X_i\rangle_{i\in I} \to \langle Y_i\rangle_{i\in I}$ and $\langle X_i\rangle_{i\in I} \to \langle Z_i\rangle_{i\in I}$, then for every $i \in I$, $Y_i \xleftarrow{\pi_{Y_i}} X_i \xrightarrow{z_i} Z_i$ in S_i, which in turn implies that $Y_i \wedge Z_i$ exists in S_i by definition of conditional meet semilattice; the reader can verify that

$$\langle Y_i\rangle_{i\in I} \wedge \langle Z_i\rangle_{i\in I} = \langle Y_i \wedge Z_i\rangle_{i\in I}.$$

The functor $\mathcal{M}$ also has the desired properties. It is clear that $\mathcal{E}(\langle \top_{S_i}\rangle_{i\in I}) \cong \{*\}$. If I is countable, then for any $(x_i)_{i\in I} \in \mathcal{E}(\langle X_i\rangle_{i\in I})$ the singleton $\{(x_i)_{i\in I}\}$ belongs to $\mathcal{B}(\langle X_i\rangle_{i\in I})$, because it can be written as a countable intersection $\bigcap_{i\in I} (p^i_{\langle X_i\rangle_{i\in I}})^{-1}(x_i)$. Finally, when $\mathcal{M}$ is applied to the product $\langle Y_i\rangle_{i\in I} \wedge \langle Z_i\rangle_{i\in I}$ and its projections, one gets

$$\mathcal{M}\langle Y_i\rangle_{i\in I} \xleftarrow{\mathcal{M}\langle \pi_{Y_i}\rangle_{i\in I}} \mathcal{M}\langle Y_i \wedge Z_i\rangle_{i\in I} \xrightarrow{\mathcal{M}\langle \pi_{Z_i}\rangle_{i\in I}} \mathcal{M}\langle Z_i\rangle_{i\in I}.$$

The map

$$\mathcal{M}\langle\pi_{Y_i}\rangle_{i\in I}\times\mathcal{M}\langle\pi_{Z_i}\rangle_{i\in I}:\mathcal{M}\langle Y_i\wedge Z_i\rangle_{i\in I}\to\mathcal{M}\langle Y_i\rangle_{i\in I}\times\mathcal{M}\langle Z_i\rangle_{i\in I}$$

is injective, because for any $(y_i)_{i\in I}\in\mathcal{E}(\langle Y_i\rangle_{i\in I})$ and $(z_i)_{i\in I}\in\mathcal{E}(\langle Z_i\rangle_{i\in I})$, the elementary properties of set operations imply that

$$\begin{aligned}
&(\mathcal{M}\langle\pi_{Y_i}\rangle_{i\in I}\times\mathcal{M}\langle\pi_{Z_i}\rangle_{i\in I})^{-1}\left((y_i)_{i\in I},(z_i)_{i\in I}\right)\\
&\quad=\mathcal{M}\langle\pi_{Y_i}\rangle_{i\in I}^{-1}\left((y_i)_{i\in I}\right)\cap\mathcal{M}\langle\pi_{Z_i}\rangle_{i\in I}^{-1}\left((z_i)_{i\in I}\right)\\
&\quad=\left\{\prod_{i\in I}\mathcal{M}_i\pi_{Y_i}^{-1}(y_i)\right\}\cap\left\{\prod_{i\in I}\mathcal{M}_i\pi_{Z_i}^{-1}(z_i)\right\}\quad\text{(by def. of }\mathcal{M})\\
&\quad=\prod_{i\in I}\left\{\mathcal{M}_i\pi_{Y_i}^{-1}(y_i)\cap\mathcal{M}_i\pi_{Z_i}^{-1}(z_i)\right\},
\end{aligned}$$

and the cardinality of each factor in the last expression is at most 1.

For each $i\in I$, we introduce a morphism of information structures $p^i:(\mathbb{S},\mathcal{M})\to(\mathbb{S}_i,\mathcal{M}_i)$ such that p_0^i maps each object or morphism $\langle A_i\rangle_{i\in I}$ to A_i, and

$$\widehat{p}^i_{\langle X_i\rangle_{i\in I}}:\prod_{i\in I}\mathcal{M}_i(X_i)\to\mathcal{M}_i(X_i)$$

is the canonical projection (which is measurable by definition of the product σ-algebra, see above). We claim that $\mathbb{S}$, with the projections p^i just introduced, is the product of $(\mathbb{S}_i,\mathcal{M}_i)_{i\in I}$ in **InfoStr**, written $\prod_{i\in I}(\mathbb{S}_i,\mathcal{M}_i)$, unique up to unique isomorphism (we also use the symbol $\times$ for finite products). In fact, given an I-cone $\{f^i:(\mathbb{R},\mathcal{F})\to(\mathbb{S}_i,\mathcal{M}_i)\}_{i\in I}$ in **InfoStr** (where I is seen as a discrete category), define $\langle f^i\rangle_{i\in I}:(\mathbb{R},\mathcal{F})\to(\mathbb{S},\mathcal{M})$ by

$$\left(\langle f^i\rangle_{i\in I}\right)_0:\mathbb{R}\to\mathbb{S}$$
$$R\mapsto\langle f^i(R)\rangle_{i\in I}$$

for any object or morphism R; for any $X\in\mathrm{Ob}\,\mathbb{R}$, the surjection

$$\widehat{\langle f^i\rangle_{i\in I}}(X):\mathcal{F}(X)\to\mathcal{M}(\langle f^i(X)\rangle_{i\in I})=\prod_{i\in I}\mathcal{M}_i(f^i(X))$$

is the map $\langle\widehat{f^i_X}\rangle_{i\in I}$ induced by the I-cone $\{\widehat{f^i_X}:\mathcal{F}(X)\to\mathcal{M}_i(f^i(X))\}_{i\in I}$ in **Sets**, in such a way that $p^i\circ\langle f^j\rangle_{j\in I}=f^i$ for all $i\in I$.

Nonempty coproducts: Given information structures $\{(\mathbb{S}_i,\mathcal{M}_i)\}_{i\in I}$, define a category $\mathbb{S}$ such that

- $\mathrm{Ob}\,\mathbb{S}=\coprod_{i\in I}\mathrm{Ob}\,\mathbb{S}_i/\sim$, where $\sim$ is the smallest equivalence relation such that $\top_{\mathbb{S}_i}\sim\top_{\mathbb{S}_j}$ for all $i,j\in I$;

- $A \to B$ in $\mathbb{S}$ if and only if $A \to B$ in $\mathbb{S}_i$ for some i.

Let $\mathcal{M} : \mathbb{S} \to$ **Sets** be the functor that coincides with $\mathcal{M}_i$ on $\mathbb{S}_i$. The pair $(\mathbb{S}, \mathcal{M})$ is an information structure: the properties in Definition 2.4 are verified locally on each $\mathbb{S}_i$.

Injections $j^i : \mathbb{S}_i \to \mathbb{S}$ are defined in the obvious way: $j^i_0(A) = A$ for $A \in \text{Ob } \mathbb{S}_i$ or $A \in \text{Hom}(\mathbb{S}_i)$, and the mappings $\widehat{j}^i_X$ are identities. If $\{f^i : (\mathbb{S}_i, \mathcal{M}_i) \to (\mathbb{R}, \mathcal{F})\}_{i \in I}$ is an I-cocone, define

$$\left(\langle f^i \rangle_{i \in I}\right)_0 : \mathbb{S} \to \mathbb{R}$$

$$A \mapsto f^i(A) \text{ if } A \in \text{Ob } \mathbb{S}_i \text{ or } A \in \text{Hom}(\mathbb{S}_i)$$

and, if $X \in \text{Ob } \mathbb{S}_i$, set $(\langle f^j \rangle_{j \in I})_X = \widehat{f^i}_X$. By construction, $\langle f^j \rangle_{j \in I} \circ j^i = f^i$. Therefore, $(\mathbb{S}, \mathcal{M})$ is the coproduct of $\{(\mathbb{S}_i, \mathcal{M}_i)\}_{i \in I}$ in **InfoStr**, denoted $\coprod_{i \in I} (\mathbb{S}_i, \mathcal{M}_i)$ (we also use $\bigsqcup$ for finite coproducts), which is unique up to unique isomorphism. $\blacksquare$ $\square$

Remark 2.2. *If $(\mathbf{S}_1, \mathcal{M}_1)$ and $(\mathbf{S}_2, \mathcal{M}_2)$ are bounded structures, their product and coproduct are bounded too. In fact, if the height of the poset $\mathbf{S}_i$ is N_i ($i = 1, 2$), then the height of $\mathbf{S}_1 \times \mathbf{S}_2$ is $N_1 + N_2$ and that of $\mathbf{S}_1 \sqcup \mathbf{S}_2$ equals $\max(N_1, N_2)$. Similarly, if both structures are finite, their product and coproduct is finite too.*

Remark 2.3. *If each measurable space $(\mathcal{E}(X), \mathcal{B}(X))$ appearing in $\mathbf{S}_1$ and $\mathbf{S}_2$ verifies that $\mathcal{E}(X)$ is second countable topological space and $\mathcal{B}(X)$ is its Borel σ-algebra, then each algebra $\mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$ on $\mathcal{E}(X_1) \times \mathcal{E}(X_2)$ equals the Borel σ-algebra on this space.*

2.4 Representations

We introduce here the notion of **representation** of an information structure in terms of classical observables (measurable functions) or quantum observables (self-adjoint operators), as a bridge between our categorical definitions and the more traditional models used in classical and quantum probability theory. The rest of the paper does not depend on this section. For simplicity, we restrict ourselves to finite information structures.

Recall that $\mathbf{Obs}_{\text{fin}}(\Omega)$ denotes the poset of finite partitions of a set Ω, ordered by the relation of refinement, and $\square$ is the "forgetful" functor from $\mathbf{Obs}_{\text{fin}}(\Omega)$ into **Sets** introduced in example 2.2, that maps each partition $\{A_1, ..., A_n\}$ to the set $\{A_1, ..., A_n\}$ and each arrow of refinement to a surjection.

Definition 2.17 (representation)**.** *A classical representation of a **finite information structure** $(\mathbf{S}, \mathscr{E})$ is a pair (Ω, ρ), where Ω is a set and $\rho = (\rho_0, \widehat{\rho}) : (\mathbf{S}, \mathscr{E}) \to (\text{Obs}_{\text{fin}}(\Omega), \square)$ is a morphism of information structures such that $\widehat{\rho} : \mathscr{E} \to \rho_0^* \square$ is a natural isomorphism (i.e. the components $\widehat{\rho}_X : \mathscr{E}(X) \to \square\rho_0(X)$ are bijections, natural in X).*

If (Ω, ρ) is a classical representation of $\mathbf{S}$, each observable X in $\mathbf{S}$ can be associated with a unique function $\widetilde{X} : \Omega \to \mathscr{E}_X$, in such a way that $\rho_0(X)$ is the partition induced by $\widetilde{X}$ and $\widehat{\rho}_X(x) = \widetilde{X}^{-1}(x)$. Since ρ_0 is a morphism of conditional *meet* semilattices, for any $X, Y \in \operatorname{Ob}\mathbf{S}$ the joint $(\widetilde{X}, \widetilde{Y}) : \Omega \to \mathscr{E}_X \times \mathscr{E}_Y$ is equivalent to $\widetilde{X \wedge Y} : \Omega \to \mathscr{E}_{XY}$, in the sense that both induce the same partition of Ω.

The next proposition points to a close link between representations and $\lim \mathscr{E}$. Recall that the limit of the functor $\mathscr{E} : \mathbf{S} \to \mathbf{Sets}$ is defined as

$$\lim \mathscr{E} := \operatorname{Hom}_{[\mathbf{S}, \mathbf{Sets}]}(*, \mathscr{E}),$$

where $[\mathbf{S}, \mathbf{Sets}]$ is the category of covariant functors from $\mathbf{S}$ to $\mathbf{Sets}$, and $*$ is the functor that associates to each object a one-point set; equivalently

$$\lim \mathscr{E} \cong \left\{ (s_Z)_{Z \in \operatorname{Ob}\mathbf{S}} \in \prod_{Z \in \operatorname{Ob}\mathbf{S}} \mathscr{E}(Z) : \mathscr{E}_{\pi_{YX}}(s_X) = s_Y \text{ for all } \pi_{YX} : X \to Y \right\},$$

where s_Z denotes $\varphi(*)$ for any $\varphi \in \operatorname{Hom}_{[\mathbf{S}, \mathbf{Sets}]}(*, \mathscr{E})$. The requirements imposed on $(s_Z)_{Z \in \operatorname{Ob}\mathbf{S}}$ in (23) are referred hereafter as **compatibility conditions**. We denote the restriction of each projection $\pi_{\mathscr{E}(X)} : \prod_{Z \in \operatorname{Ob}\mathbf{S}} \mathscr{E}(Z) \to \mathscr{E}(X)$ to $\lim \mathscr{E}$ by the same symbol. We interpret the limit as all possible combinations of compatible outcomes.

Proposition 2.5. *If $(\mathbf{S}, \mathscr{E})$ has a classical representation $(\Omega, \rho, \widehat{\rho})$, then for any $X \in \operatorname{Ob}\mathbf{S}$ and any $x \in \mathscr{E}(X)$, there exists an element $s(x) \in \lim \mathscr{E}$ such that $\pi_{\mathscr{E}(X)}(s(x)) = x$.*

Proof. For each X in $\operatorname{Ob}\mathbf{S}$, there is a surjection $\pi_X : \Omega \to \mathscr{E}(X)$ obtained as the composition of $\xi_X : \Omega \to \square\rho_0(X)$, which maps $\omega \in \Omega$ to the part that contains it, and $\widehat{\rho}_X^{-1} : \square\rho_0(X) \to \mathscr{E}(X)$. The maps $\{\pi_X\}$ define a cone over $\mathscr{E}$ i.e. given $f : Y \to Z$ in $\mathbf{S}$, one has a commutative diagram

$$\begin{tikzcd}
\Omega \arrow[r, "\xi_Y"] \arrow[d, equal] & \square\rho_0(Y) \arrow[r, "\widehat{\rho}_Y^{-1}"] \arrow[d, "\square\rho_0 f"'] & \mathscr{E}(Y) \arrow[d, "\mathscr{E} f"] \\
\Omega \arrow[r, "\xi_Z"] & \square\rho_0(Z) \arrow[r, "\widehat{\rho}_Z^{-1}"] & \mathscr{E}(Z)
\end{tikzcd}$$

the commutation of the left square comes from the definition of $\mathbf{Obs}_{\text{fin}}(\Omega)$ and $\square$, and the right square commutes because $\widehat{\rho}$ is a natural isomorphism. Therefore, there is a map $\pi : \Omega \to \lim \mathscr{E}$ and the desired section is obtained as the image under π of any $\omega \in \widehat{\rho}_X(x)$. ■ □

References

[1] Juan Pablo Vigneaux. "Information structures and their cohomology". In: *arXiv preprint arXiv:1709.07807* (2017).

[2] Pierre Baudot and Daniel Bennequin. "The homological nature of entropy". In: *Entropy* 17.5 (2015), pp. 3253–3318.

[3] Claude E Shannon. "A mathematical theory of communication". In: *The Bell system technical journal* 27.3 (1948), pp. 379–423.