BRUNO FORTE

Measures of information : the general axiomatic theory

<http://www.numdam.org/item?id=M2AN_1969__3_2_63_0>

MEASURES OF INFORMATION :
THE GENERAL AXIOMATIC THEORY

by Bruno FORTE ([1])

Summary. — *A general definition of information measure is given by means of some basic properties. Some special properties are also considered. They lead to new characterizations of Shannon's and Renyi's information measures for incomplete probability distributions.*

1. INTRODUCTION

Last year J. Kampé de Fériet and I (see [1]) pointed out the motivations for a direct definition of the measure of information making no use of the concept of probability. Of course we started restricting ourselves to the simple case of measures of information $J(A)$ given by single events A. We stated the axioms which define the measures of information, then we studied (see [2] and [3]) some special classes of information measures.

More recently we suggested a reasonable definition for the measure of information given by an experiment, treating the case of complete information distributions first (see [4]), then the case of incomplete distributions (see [5]); in both cases the localization property was assumed to be exhibited by our general measures of information. That property, as known, is closely related to the classical Faddeev axiom (see [6], [7], [8]); therefore no wonder if we did not include in our definition some interesting information measures, Rény's measure, for instance, which exhibit no localization property.

Hence it seems suitable to look for a general definition of information measure which covers all the cases and includes all the classical information measures for probability distributions. In the present paper we will go through this problem.

Starting from a reasonably general definition for the information measure we will consider some classes of information measures in order to get also new characterizations for Shannon's and Rény's entropies.

(1) Bruno Forte, Istituto di Matematica, Università di Pavia, 27 100, Pavia, Italy.

2. INFORMATION MEASURES, INFORMATION SPACES

Let Ω (sure event) be a set of elements ω (elementary events). Let S be a non empty class of subsets A of Ω. Let ε be a class of partitions π_A (experiments) of subsets $A \in S$ into a finite number of non empty subsets $A_i \in S$, $A_i \neq \emptyset$. This class is non empty, for it will contain at least all the non empty sets in S.

Definition 1. If κ_a is a finite or infinite class of partitions in ε, we say that the partitions of the class κ_a are algebraically independant if

$$\bigcap_{r=1}^{r=n} \tilde{A}_{r,h_r} \neq \emptyset$$

for every finite class $\{ \pi_{A_r} = (A_{r,1}, A_{r,2}, ..., A_{r,m_r}) : r = 1, 2, ..., n$ e of distinct partitions in κ_a and for every choice of each non empty set $\tilde{A}_r (\neq \Omega)$ the algebra $\mathcal{A}_r$ generated by $(A_{r,1}, A_{r,2}, ..., A_{r,m_r})$.

Definition 2. If κ_a is an infinite class of partitions in ε, we say that the partitions of the class κ_a are algebraically σ-independent if

$$\bigcap_{r=1}^{r=+\infty} \tilde{A}_{r,h_r} \neq \emptyset$$

for every sequence $\{ \pi_{A_r} = (A_{r,1} A_{r,2} ..., A_{r,m_r}) : r = 1, 2, ..., \}$ of distinct partitions in κ_a and for every choice of each non empty set $\tilde{A}_r (\neq \Omega)$ in the algebra generated by $(A_{r,1}, A_{r,2}, ..., A_{r,m_r})$.

Definition 3. If π_A and π_B are two partitions in ε, the operations $\cup$ (union) $\cap$ (intersection) are defined by

$$\pi_A \cup \pi_B = \{ C : \text{either } C \in \pi_A \text{ or } C \in \pi_B \} = \pi_{A \cap B}$$

for every $\pi_A \in \varepsilon$, $\pi_B \in \varepsilon$, $A \cap B = \emptyset$,

$$\pi_A \cap \pi_B = \{ C : C = A_i \cap B_j, A_i \in \pi_A, B_j \in \pi_B \} = \pi_{A \cup B}.$$

These two operations are commutative and associative; furthermore each of them is distributive with respect to the other.

In view of the next definition let us consider the set $I(\pi_A)$ defined for every partition $\pi_A \equiv (A_1, A_2, ..., A_m) \in \varepsilon$ as follows

$$I(\pi_A) = \emptyset \quad \text{for} \quad m = 1$$

$$I(\pi_A) = \Omega \quad \text{for} \quad m > 1,$$

and then for each pair of partitions

$$\pi_A \equiv (A_1, A_2, ..., A_m) \quad \text{and} \quad \pi_B \equiv (B_1, B_2, ..., B_n)$$

let us call $D(\pi_A, \pi_B)$ the set

$$(A \cup B - A \cap B) \cap [I(\pi_A) \cup I(\pi_B)]$$

with $A = \bigcup_{i=1}^{i=m} A_i$, $B = \bigcup_{i=1}^{i=n} B_i$. It is clear that for $m = n = 1$ we have always $D(\pi_A, \pi_B) = \varnothing$ and that for $mn \neq 1$ $D(\pi_A, \pi_B) = \varnothing$ implies $A = B$.

Definition 4. If π_A and π_B are two partitions such that

$$D(\pi_A, \pi_B) = \varnothing$$

we shall say that π_B is a refinement of π_A and we shall write

$$\pi_A < \pi_B$$

if

$$\forall\, B_i \in \pi_B \,\exists\, A_j \in \pi_A \ni B_i \subset A_j$$

The relation $<$ between partitions is always reflexive; besides $\pi_A < \pi_B$, $\pi_B < \pi_C$, $D(\pi_A, \pi_C) = \varnothing$ imply

$$\pi_A < \pi_C.$$

Definition 5. An information measurable space is a set Ω a non empty class S of subsets of Ω, a class ε of partitions of sets $A \in S$ into a finite number of non empty sets $A_i \in S$, a collection κ^* of classes κ_{in} of algebraically (σ-) independant partitions in ε, such that all intersections of a finite (infinite) number of distinct partitions in κ_{in} belong to ε.

Note that κ^* is the collection of the classes of experiments which are assumed a priori to be (σ —) independent with respect to information. It is obvious that this collection may be empty.

We shall indicate a particular information measurable space with $(\Omega, S, \varepsilon, \kappa^*)$.

Definition 6 (Information measure). Let $(\Omega, S, \varepsilon, \kappa^*)$ be an information measurable space. An information measure is an extended real valued non negative function H defined on ε with the following properties :

a) (*monotonicity*) $\pi_A < \pi_B$ implies $H(\pi_A) \leqslant H(\pi_B)$.

b) (*additivity*) if $\kappa_{in} \in \kappa^*$ is a class of independent partitions, then

$$H\left(\bigcap_{r=1}^{r=n} \pi_{A_r}\right) = \sum_{r=1}^{r=n} H(\pi_{A_r}),$$

for every finite class $\{\, \pi_{A_r} : r = 1, 2, ..., n \,\}$ of distinct partitions in κ_{in};

if $\kappa_{in} \in \kappa^*$ is a class of σ-independent partitions, then

$$H\left(\bigcap_{r=1}^{r=+\infty} \pi_{A_r}\right) = \sum_{r=1}^{r=+\infty} H(\pi_{A_r})$$

for every sequence $\{ \pi_{A_r} : r = 1, 2, ... \}$ of distinct partitions in κ_{in}.

Definition 7 (Information space). An information space is an information measurable space $(\Omega, S, \varepsilon, \kappa^*)$ and an information measure H on ε.

According to the previous notation we shall indicate a particular information space with $(\Omega, S, \varepsilon, \kappa^*, H)$.

In what follows ε_n will represent the collection of the partitions in ε which consist of n sets. It is clear that $\varepsilon_1 \subset S$, the validity of

$$\sum_{n=1}^{n=+\infty} \varepsilon_n = \varepsilon$$

is also obvious.

3. EXAMPLES

We shall illustrate here the definition of information space by means of some examples.

EXAMPLE 1. Let $(\Omega, S, \varepsilon, \kappa^*)$ be any information measurable space. Denote by N the set of positive integers, then consider the function $n : \varepsilon \to N$, which for every $\pi_A \in \varepsilon$ assigns the number of sets A_i in π_A. Denote now by $\bar{R}^+$ the set of the extended real non negative numbers and consider a function $\varphi : n(\varepsilon) \to \bar{R}^+$ with the following properties

$a')$ φ is non decreasing.

$b')$ if $\{ \pi_{A_r} : r = 1, 2, ..., m \}$ is a finite class of distinct partitions in $\kappa_{in} \in \kappa^*$ then

$$\varphi\left[n\left(\bigcap_{r=1}^{r=m} \pi_{A_r} \right) \right] = \varphi[n(\pi_{A_1})n(\pi_{A_2}) ... n(\pi_{A_m})] = \sum_{r=1}^{r=m} \varphi[n(\pi_{A_r})],$$

if $\{ \pi_{A_r} : r = 1, 2, ... \}$ is a sequence of distinct partitions in $\kappa_{in} \in \kappa^*$ then

$$\varphi\left[n\left(\bigcap_{r=1}^{r=+\infty} \pi_{A_r} \right) \right] = \varphi[n(\pi_{A_1})n(\pi_{A_2}) ...] = \sum_{r=1}^{r=+\infty} \varphi[n(\pi_{A_r})].$$

Set $H(\pi_A) = \varphi[n(\pi_A)]$. It is easy to see that this space $(\Omega, S, \varepsilon, \kappa^*, H)$ is an information space. We can choose in particular $H(\pi_A) = c \log n(\pi_A)$; in this case no restriction has to be imposed to κ^* in order that (a') and (b') be verified.

EXAMPLE 2. Let Ω be an infinite set of elements or points ω. Let S be any non empty class of non empty subsets A of Ω. Let ε be any class of partitions of sets $A \in S$ into a finite number of subsets in S. Let κ^* be a collection of classes κ_{in} of $(\sigma \longrightarrow)$ independent partitions in ε, such that the intersection of a finite class (sequence) of distinct partitions in κ_{in} has an infinite number of

points ω. For every $A \in S$ denote by $n(A)$ the number of points in A. Then consider the function $H : \varepsilon \to \bar{R}^+$ defined by

$$H(\pi_A) = \sum_{i=1}^{i=m} [n(A_i)]^{-1}.$$

It is easy to verify that $(\Omega, S, \varepsilon, \kappa^*, H)$ is an information space.

EXAMPLE 3. Let X be a metric space. Let Ω be an open set of X. Let us fix a point $\omega_0 \in \Omega$ and let us denote by $S(\omega_0; r) = \{ \omega \in X; d(\omega, \omega_0) < r \}$ the open sphere with centre ω_0 and radius r. Let S be the class of the spheres $S(\omega_0; r)$ which are contained in Ω. If ε is a class of partitions of sets $A \in S$ into a finite number of subsets in ε, then it is easy to see that ε is reduced to ε_1. It is also evident that because of the definition of S the collection κ^* is empty. Consider now the function $H : \varepsilon \to R$ defined by

$$H[S(\omega_0; r)] = h(r) \qquad \text{for every} \qquad S(\omega_0; r) \in \varepsilon \equiv \varepsilon_1,$$

where $h(r)$ is any extended real valued non negative and non increasing function. This space $(\Omega, S, \varepsilon, \kappa^*, H)$ is an information space (see [9] and [10]).

EXAMPLE 4. Let us start again from a metric space X. Let Ω be any non empty subset of X, S a class of non empty subsets of Ω. If $A \in S$ let us denote by $\delta(A)$ its diameter. As usual let ε be a class of partitions of sets $A \in S$ into a finite number of subsets $A_i \in S$. Let finally κ^* be a collection of classes κ_{in} of $[\sigma -]$ independent partitions in ε. Consider the function $H : \varepsilon \to \bar{R}^+$ defined by

$$H(\pi_A) = [\operatorname*{Max}_{A_i \in \pi_A} \delta(A_i)]^{-1} \qquad \text{for every } \pi_A \in \varepsilon \; ;$$

with a suitable choice of κ^*, which will be eventually empty, the space $(\Omega, S, \varepsilon, \kappa^*, H)$ is an information space.

EXAMPLE 5. With the same notations as in example 4, for every partition π_A in a class ε of partitions of sets A in a metric space such that $\delta(\Omega) < + \infty$, let us consider the function H defined by

$$H(\pi_A) = \frac{1}{\delta(\Omega)} \left[\frac{\delta(\Omega)}{\delta(A)} \right]^m$$

where m is the number of disjoint sets A_i in π_A, $\pi_A = (A_1, A_2, ..., A_m)$. Then with a suitable choice of κ^* the space $(\Omega, S, \varepsilon, \kappa^*, H)$ is an information space.

EXAMPLE 6. We start now from a measure space (X, S, μ). Suppose furthermore that X be a metric space. Let Ω be again a non empty set in S, S a classe of non empty sets $A \in S$, ε a collection of partitions of sets $A \in S$ into a finite number of sets in S. For every $\pi_A = (A_1, A_2, ..., A_m)$ consider the function H defined by

$$H(\pi_A) = \frac{\sum_{i=1}^{i=m} \delta(A_i)}{\mu(A)\delta(A)}.$$

Then with a suitable choice of κ^*, $(\Omega, S, \varepsilon, \kappa^*, H)$ is an information space.

EXAMPLE 7. This example and the next one are related to a probability space (Ω, S, P), where Ω is a set (sure event) of elements ω (elementary events), S is a Boolean σ-algebra of subsets (events) of Ω, P is a measure on S such that $P(\Omega) = 1$. Now let S be a collection of non empty sets in S, ε a collection of partitions of sets A in S into a finite number of sets in S. Among the classes of algebraically and also stochastically $(\sigma-)$ independent partitions select the classes κ_{in} to form κ^*, Set

$$H(\pi_A) = - \frac{\sum_{i=1}^{i=m} P(A_i) \log_a P(A_i)}{\sum_{i=1}^{i=m} P(A_i)} \cdots,$$

where $\pi_A = (A_1, A_2, ..., A_m)$ is in ε. This function $H : \varepsilon \to \bar{R}^+$ is an information measure (*Shannon measure*) and $(\Omega, S, \varepsilon, \kappa^*, H)$ is an information space (*Shannon's information space*).

EXAMPLE 8. In the same situation and with the notations of example 7 let us write

$$H(\pi_A) = \frac{1}{1 - \alpha} \log_a \frac{\sum_{i=1}^{i=m} [P(A_i)]^\alpha}{\sum_{i=1}^{i=m} P(A_i)} \quad , \qquad \text{with } \alpha \neq 1.$$

This function $H : \varepsilon \to \bar{R}^+$ defines an information measure (*Rényi's measure*) on $(\Omega, S, \varepsilon, \kappa^*)$ and $(\Omega, S, \varepsilon, \kappa^*, H)$ is then an information space (*Rényi's information space*).

4. THE RESTRICTION OF H TO ε_1

Let us observe previously that ε_1 is just a class of non empty sets in S. Then from the above definition 6 we have that the restriction of H to ε_1 is a set function H_1 defined on ε_1 with the following properties

I) $\varepsilon_1 \overset{H_1}{\to} \bar{R}^+$,

II) $A \in \varepsilon_1$, $B \in \varepsilon_1$, $B \subset A \Rightarrow H_1(B) \geqslant H_1(A)$,

III) $H_1\left(\bigcap_{i \in I} B_i \right) = \sum_{i \in I} H_1(B_i)$

for every non empty finite (independence) or countable (σ-independence) index set I such that $B_i \in \kappa_{in}$.

Hence H_1 is a measure of information J defined on the class of sets (events) ε_1 according to the definition of J given in [1] and [10].

Going back to the examples of Sect. 3, its is easy to recognize that they lead to the following measures of information J :

1) $$J(A) = \varphi(1)$$

this means that all events $A \in \varepsilon_1$ give the same amount of information.

2) $$J(A) = \frac{1}{n(A)}$$

the information given by A is inversely proportional to the number of points in A.

3) $A = S(\omega_0; r)$ and $J(A) = h(r)$.

4) }
5) } $$J(A) = \frac{1}{\delta(A)} \cdot$$

6) $$J(A) = \frac{1}{\mu(A)} \cdot$$

7) }
8) } $$J(A) = - \log_a P(A).$$

Thus every information measure H assigns a uniquely determined information function J on ε_1. In a next section we shall deal with the converse problem that is the problem of the characterization of the information measure which can be evaluated starting from an information measure J defined on ε_1. This can be done for instance with (2), (4), (5), (7) and (8) but not with the examples (1) and (6).

5. INEQUALITIES

Referring to the operations $\cup$ and $\cap$ between events and partitions (experiments) the monotonicity of H implies a set of inequalities which have a certain degree of interest; since they can be deduced trivially from 2-a and 4-II we shall omit the proofs.

Proposition 1. For each pair of events $A \in \varepsilon_1$ and $B \in \varepsilon_1$ such that $A \cup B \in \varepsilon_1$ we have

$$J(A \cup B) \leqslant \text{Inf} \, [J(A), J(B)].$$

Proposition 2. If $A \in \varepsilon_1$, $B \in \varepsilon_1$ and $A \cap B \in \varepsilon_1$ then we have

$$J(A \cap B) \geqslant \text{Sup} \, [J(A), J(B)].$$

Proposition 3. If $\Omega \in \varepsilon_1$ then $J(\Omega) \leqslant \underset{A \in \varepsilon_1}{\text{Inf}} \, J(A)$

If we include in ε_1 the empty set $\varnothing$ we can assign to its amount of information $J(\varnothing)$ any value such that $J(\varnothing) \geqslant \underset{A \in \varepsilon_1}{\text{Sup}} \, J(A)$, according to the property 4-II.

Henceforth we shall denote by $\{ A \}$ the partitions π_A which consist of a single event A. As it concerns partitions the following general inequalities hold :

Proposition 4. For every pair (π_A, π_B) of partitions in ε such that $A \cap B = \varnothing$, $\pi_A \cup \pi_B \in \varepsilon$, $\pi_A \cup \{ B \} \in \varepsilon$ and $\pi_B \cup \{ A \} \in \varepsilon$ we have

$$H(\pi_A \cup \pi_B) \geqslant \text{Sup} \, [H(\pi_A \cup \{ B \}), H(\pi_B \cup \{ A \})].$$

Proposition 5. If $\pi_A \in \varepsilon$, $\pi_B \in \varepsilon$ and $\pi_A \cap \pi_B \in \varepsilon$, $\pi_A \cap \{ B \} \in \varepsilon$ and $\pi_B \cap \{ A \} \in \varepsilon$, then we have

$$H(\pi_A \cap \pi_B) \geqslant \text{Sup} \, [H(\pi_A \cap \{ B \}), H(\pi_B \cap \{ A \})].$$

6. LOCAL MEASURES OF INFORMATION

We begin with the definition of what is the localization property.

Definition 8. We shall say that a measure of information is a local information measure or that it exhibit the localization property if

$$H(\pi_A \cup \pi_B) - H(\pi_A \cup \{ B \}) = H(\{ A \} \cup \pi_B) - H(\{ A \} \cup \{ B \})$$

holds for every $\pi_A \in \varepsilon$ and $\pi_B \in \varepsilon$ such that $A \cap B = \varnothing$ and $\pi_A \cup \pi_B \in \varepsilon$, $\pi_A \cup \{ B \} \in \varepsilon$, $\{ A \} \cup \pi_B \in \varepsilon$, $\{ A \} \cup \{ B \} \in \varepsilon$.

In particular if for every $\pi_A = (A_1, A_2, ..., A_n)$

$$\{ A_i \} \in \varepsilon \, (i = 1, 2, ..., n) \quad , \quad \pi'_A < \pi_A \Rightarrow \pi'_A \in \varepsilon$$

and if

$$H(\pi_A) - H(\pi_{A - A_{n-1}} \cup {}_{A_n} \cup \{ A_{n-1} \cup A_n \}) = H(A_1, A_2, ..., A_n) -$$

$$- H(A_1, A_2, ..., A_{n-1} \cup A_n) = \Phi \left[H \left(\bigcup_{i=1}^{i=n-2} A_i, H(A_{n-1}), H(A_n) \right) \right]$$

where Φ is a real valued non negative function with domain the set

$$\Gamma_3 = \{ x, y, z : x = H(\{ B \}), y = H(\{ C \}), z = H(\{ D \}),$$

$$\{ B \} \in \varepsilon, \{ C \} \in \varepsilon, \{ D \} \in \varepsilon, B \cap C = \varnothing, C \cap D = \varnothing, D \cap B = \varnothing \},$$

then the information measure H is a local measure of information (see [5]).

This special property is what has been called *branching principle* (see [7]).

The branching principle is for instance the special property exhibited by Shannon's measure (see [7] and [8]).

Going back to the examples we have considered in Sect. 3 it is a simple matter to verify that the information measure of example 2 is also a local measure of information; on the contrary Renyi's measure (example 8) and the measures of examples 1, 4, 5 and 6 do not exhibit in general the localization property.

7. IDEMPOTENT INFORMATION MEASURES

The following property as well as the localization property will come useful for the classification of different information measure.

Definition 9. An information measure H on an information measurable space $(\Omega, S, \varepsilon, \kappa^*)$ is called an *idempotent measure* if from

$$\pi_{A_\alpha} \in \varepsilon \; \forall \; \alpha \in I, \forall \; \alpha, \beta \in I \quad \alpha \neq \beta \quad A_\alpha \cap A_\beta = \varnothing, \quad \bigcup_{\alpha \in I} \pi_{A_\alpha} \in \varepsilon$$

and $\forall \; \alpha, \beta \in I$

$$H(\pi_{A_\alpha}) = H(\pi_{A_\beta}) = \bar{H},$$

where I is any non empty index set, it follows

$$H\left(\bigcup_{\alpha \in I} \pi_{A_\alpha}\right) = H(\pi_{A_\alpha}) = \bar{H}.$$

Thus if H is an idempotent measure of information, for all partitions

$$\pi_A = (A_1, A_2, ..., A_m) \in \varepsilon$$

such that $\{A_i\} \in \varepsilon$ and $J(A_i) = \bar{J} (i = 1, 2, ..., m)$ we have

$$H(\pi_A) = J(A_i) = \bar{J}.$$

It is easy to find examples which prove that the converse is not true.

Shannon's and Renyi's entropies are both idempotent measures of information, as well as the measure of example 4, while the measures of examples 1, 2, 5 and 6 are not idempotent.

Proposition 6. If for every $\pi_{A_\alpha} \in \varepsilon (\alpha \in I)$ such that

$$\forall \; \alpha \in I, \beta \in I, \alpha \neq \beta, A_\alpha \cap A_\beta = \varnothing \qquad \text{and} \qquad \bigcup_{\alpha \in I} \pi_{A_\alpha} \in \varepsilon$$

the following inequalities hold

$$\operatorname*{Inf}_{\alpha \in I} H(\pi_{A_\alpha}) \leqslant H\left(\bigcup_{\alpha \in I} \pi_{A_\alpha}\right) \leqslant \operatorname*{Sup}_{\alpha \in I} H(\pi_{A_\alpha})$$

then the measure H is an idempotent measure of information.

Proof : for $H(\pi_{A_\alpha}) = \overline{H}$ the above inequalities imply

$$H\left(\bigcup_{\alpha \in I} \pi_{A_\alpha}\right) = \overline{H}.$$

It is also easy to find examples which prove that the converse is not true, that is from the idempotence of H alone one cannot derive the above inequalities.

8. SET COMPOSITIVE INFORMATION MEASURES

Let us consider the restriction of a measure of information to ε_1 and let us devote our attention to the operation $\cup$ between the events in ε_1. Given two events A and B in ε_1 such that $A \cap B = \varnothing$, suppose that their union $A \cup B$ be in ε_1, then by proposition 1 we know that

$$J(A \cup B) \leqslant \text{Inf } [J(A), J(B)];$$

in the most of cases we cannot say anything else about the information given by the union of A and B, but in some special cases, that is for some special information spaces $(\Omega, S, \varepsilon, \kappa^*, H)$, the information given by the union of A and B for every $A \in \varepsilon_1$ and $B \in \varepsilon_1$ such that $A \cup B \in \varepsilon_1$, $A \cap B = \varnothing$, may be completely determined by the information $J(A)$ given by A and the information $J(B)$ given by B. This is the case, for example, of Shannon's and Rényi's entropies as well as the case of the measures of information we have considered in the examples 2 and 6. On this ground the following definition seems to be quite justified.

Definition 10. An information measure H on an information measurable space $(\Omega, S, \varepsilon, \kappa^*)$ is called : *set compositive measure* of information if one can find a real valued non negative function F with domain
$\Gamma_2 = \{ x, y : x = J(A), y = J(B), A \in \varepsilon_1, B \in \varepsilon_1, A \cup B \in \varepsilon_1, A \cap B = \varnothing \}$
so that

$$J(A \cup B) = F[J(A), J(B)]$$

holds for every $A \in \varepsilon_1$ and $B \in \varepsilon_1$ such that $A \cup B \in \varepsilon_1$, $A \cap B = \varnothing$.

The function F will be called *set composition law* for the restriction of H to ε_1.

As regards the examples of information measures in Sect. 3 we see that the measures of examples 4 and 5 are generally not set compositive, whereas the measures of examples 1, 2, 6, 7 and 8 are set compositive with the following composition laws.

Example 1 : $F(x, y) = \varphi(1) = \text{const.}$

Examples 2 and 6 : $F(x, y) = \dfrac{1}{\dfrac{1}{x} + \dfrac{1}{y}}$ (hyperbolic law)

Examples 7 and 8 : $F(x, y) = - \underset{c>0}{c} \log [e^{-x/c} + e^{-y/c}]$, with $c = \log_a e$, we shall call this last law Shannon's set composition law, and we shall say that an information measure possesses a *shannonian* restriction if it exhibits a shannonian composition law.

It does not make sense to look for a composition law in the case of example 3; in fact for that particular information measurable space it does not exist any pair (A, B) of events in ε_1 such that $A \cap B = \emptyset$.

It will turn out useful to list the properties which follow (see [1]) for any composition law F from the equality

$$J(A \cup B) = F[J(A), J(B)].$$

They are :

b_1) $\forall\, (x, y) \in \Gamma_2 : 0 \leqslant F(x, y) \leqslant \operatorname{Inf}(x, y),$

b_2) *(symmetry)* $\forall\, (x, y) \in \Gamma_2 : (y, x) \in \Gamma_2$ and $F(y, x) = F(x, y),$

b_3) *(monotonicity)* $\forall\, (x, y'), \in \Gamma_2, \forall\, (x, y'') \in \Gamma_2$ such that

$$y' = J(B') \quad , \quad y'' = J(B'')$$

with $B' \in \varepsilon_1, B'' \in \varepsilon_1, B' \supset B''$:

$$F(x, y') \leqslant F(x, y''),$$

b_4) *(associativity)* : $F[x, F(y, z)] = F[F(x, y), z]$

$\forall\, (x, y, z) \in \bar{\Gamma}_3 = \{ (x, y, z) : (x, y) \in \Gamma_2 \;,\; (y, z) \in \Gamma_2 \;,\; (z, x) \in \Gamma_2 \},$

b_5) if we add to ε_1 the empty set $\emptyset$, and if we set $e = J(\emptyset)$, then for every $x \in J(\varepsilon_1)$

$$F(x, e) = x,$$

b_6) *(consistency* between the two operations F and $+$) Suppose A, B, C different sets in ε_1 such that $\{ A \}$ and each non empty set of the algebra generated by B and C belong as partitions to the same class κ_{in}, respectively,

then

α) if $B \cap C \neq \emptyset$:

$$J(A) + F[F[J(B \cap C), J(B \cap (\Omega - C))], J((\Omega - B) \cap C)] = F[F[J(A)$$

$$+ J(B \cap C), J(A) + J(B \cap (\Omega - C))], J(A) + J((\Omega - B) \cap C)],$$

β) if $B \cap C = \varnothing$:

$$J(A) + F[J(B), J(C)] = F[J(A) + J(B), J(A) + J(C)]$$

has to be satisfied.

Setting $x = J(A)$, $y = J(B)$ and $z = J(C)$ the last equality assumes the following form

b_6') $x + F(y, z) = F(x + y, x + z)$ (*consistency equation*).

9. π COMPOSITIVE INFORMATION MEASURES

Let us consider now the operation $\cup$ between partitions of disjoint events. By proposition 4 we know that

$$H(\pi_A \cup \pi_B) \geqslant \text{Sup} \, [H(\pi_A \cup \{ B \}), H(\pi_B \cup \{ A \})]$$

whenever $\pi_A \cup \pi_B$, $\pi_A \cup \{ B \}$, $\pi_B \cup \{ A \}$ as well as π_A and π_B belong to ε. For the most of information spaces we cannot say anything else about the information given by the union of π_A and π_B, but for some particular information spaces it might happen that the information given by the union of π_A and π_B is completely determined by the informations given by $A \cup B$, A, B, π_A, π_B. This is for example the case of Shannon's and Rényi's measures as well as the case of the measures we have considered in the examples 2 and 4, whereas generally the remaining measures of information in those examples do not have such property. In particular this property makes no sense in the case of example 3 because in that case ε reduces to ε_1.

Therefore we give the following definition.

Definition 11. An information measure H on an information measurable space $(\Omega, \mathcal{S}, \varepsilon, \kappa^*)$ is called π *compositive measure* of information if one can find a real valued non negative function Ψ with domain

$\Gamma_5 = \{ x, y, z, u, v : x = J(A \cup B), y = J(A), z = J(B), u = H(\pi_A),$

$\qquad v = H(\pi_B), A \in \varepsilon_1, B \in \varepsilon_1, A \cup B \in \varepsilon_1, A \cap B = \varnothing, \pi_A \in \varepsilon,$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad \pi_B \in \varepsilon, \pi_A \cup \pi_B \in \varepsilon \}$

so that

$$H(\pi_A \cup \pi_B) = \Psi[J(A \cup B), J(A), J(B), H(\pi_A), H(\pi_B)]$$

holds for every $A \in \varepsilon_1$, $B \in \varepsilon_1$, $\pi_A \in \varepsilon$, $\pi_B \in \varepsilon$, such that $A \cup B \in \varepsilon_1$, $A \cap B = \varnothing$, $\pi_A \cup \pi_B \in \varepsilon$.

The function Ψ will be called the π *composition law* of H. As we told, Shannon's measure, Rényi's measure and the measure considered in examples 2 and 4 are π compositive measures. They exhibit the following composition laws :

Example 2 : $\Psi(x, y, z, u, v) = u + v$

Example 4 : $\Psi(x, y, z, u, v) = \text{Sup}\,(u, v)$

Example 7 : (Shannon's composition law)

$$\Psi(x, y, z, u, v) = \frac{u\,e^{-y/c} + v\,e^{-z/c}}{e^{-x/c}}$$

Example 8 : (Rényi's composition law)

$$\Psi(x, y, z, u, v) = \frac{c}{1-\alpha}\log\frac{e^{(1-\alpha)u/c-y/c} + e^{(1-\alpha)u/c-z/c}}{e^{-x/c}}\,,\; c = \log_a e.$$

Thus we have examples of information measures wich have no composition laws at all, examples of information measures which are set compositive but not π compositive, conversely we have examples of information measures which are π compositive and not set compositive, finally we know two measures of information which are both set compositive and π compositive (Shannon's measure and Rényi's measure).

Guided by what has been done for a set composition law F, one can easily relalize that the equality

$$H(\pi_A \cup \pi_B) = \Psi[J(A \cup B), J(A), J(B), H(\pi_A), H(\pi_B)]$$

implies the following properties for the π composition law Ψ :

c_1) $\forall\,(x, y, z, u, v) \in \Gamma_5 : \Psi(x, y, z, u, v) \geqslant x,$

c_2)   (*symmetry*) $\forall\,(x, y, z, u, v) \in \Gamma_5 : (x, z, y, v, u) \in \Gamma_5$ and

$$\Psi(x, z, y, v, u) = \Psi(x, y, z, u, v),$$

c_3)   (*monotonicity*) $\forall\,(x, y, z, u, v') \in \Gamma_5, \forall\,(x, y, z, u, v'') \in \Gamma_5$ such that

$$v' = H(\pi'_B) \qquad v'' = H(\pi''_B)$$

with

$$\pi'_B \in \varepsilon \quad,\quad \pi''_B \in \varepsilon \quad,\quad \pi'_B < \pi''_B :$$

$$\Psi(x, y, z, u, v') \leqslant \Psi(x, y, z, u, v''),$$

c_4)   (*associativity*) $\forall\,\pi_A, \pi_B, \mathrm{B}_C \in \varepsilon$ such that

$$\{A\} \in \varepsilon_1, \{B\} \in \varepsilon_1, \{C\} \in \varepsilon_1 \quad,\quad A \cap B = \varnothing,$$

$$B \cap C = \varnothing \quad,\quad C \cap A = \varnothing,$$

$\{A \cup B\} \in \varepsilon_1, \{B \cup C\} \in \varepsilon_1, \{C \cup A\} \in \varepsilon_1, \{A \cup B \cup C\} \in \varepsilon_1, \pi_A \cup \pi_B \in \varepsilon,$

$\pi_B \cup \pi_C \in \varepsilon, \pi_C \cup \pi_A \in \varepsilon, \pi_A \cup \pi_B \cup \pi_C \in \varepsilon :$

$\Psi[J(A \cup B \cup C), J(A), J(B \cup C), H(\pi_A),$

$\quad \Psi[J(B \cup C), J(B), J(C), H(\pi_B), H(\pi_C)] = \Psi[J(A \cup B \cup C), J(A \cup B), J(C),$

$\qquad\qquad \Psi[J(A \cup B), J(A), J(B), H(\pi_A), H(\pi_B)], H(\pi_C)]$

c_5)   if one adds the empty set $\emptyset$ to ε_1, then for every $\pi_A \in \varepsilon$ such that $\{ A \} \in \varepsilon_1$, setting $x = J(A), u = H(\pi_A)$, it must be

$$\Psi(x, x, e, u, e) = u,$$

with $e = J(\emptyset)$,

c_6)   (*consistency* between the two operations Ψ and $+$) define $\pi_B \cup \pi_C$ as follows

$$\pi_B \cup \pi_C = [\pi_B \cap \pi_C] \cup [\pi_B \cap (\Omega - C)] \cup [(\Omega - B) \cap \pi_C],$$

then consider the identity

$$\pi_A \cap [(\pi_B \cap \pi_C) \cup (\pi_B \cap (\Omega - C)) \cup ((\Omega - B) \cap \pi_C)]$$
$$= [\pi_A \cap \pi_B \cap \pi_C] \cup [\pi_A \cap \pi_B \cap (\Omega - C)] \cup [\pi_A \cap (\Omega - B) \cap \pi_C].$$

Suppose now :

i) A, B, C are different sets in ε_1,

ii) $\{ A \}$ and each non empty set of the algebra generated by B and C are as partitions in the same κ_{in}, respectively,

iii) π_A and each union or intersection of π_B, π_C are in the same κ_{in}, respectively,

iv) π_A and each non empty intersection of π_B and π_C with sets of the algebra generated by B and C are in the same κ_{in}, respectively.

Then

α') if $B \cap C \neq \emptyset$:

$$H(\pi_A) + \Psi \{ J(B \cup C), J(B), J((\Omega - B) \cap C), \Psi[J(B), J(B \cap C),$$
$$J(B \cap (\Omega - C)), H(\pi_B \cap \pi_C), H(\pi_B \cap (\Omega - C))], H((\Omega - B) \cap \pi_C) \}$$
$$= \Psi \{ J(A) + J(B \cup C), J(A) + J(B), J(A) + J((\Omega - B) \cap C), \Psi[J(A)$$
$$+ J(B), J(A) + J(B \cap C), J(A) + J(B \cap (\Omega - C)), H(\pi_A)$$
$$+ H(\pi_B \cap \pi_C), H(\pi_A) + H(\pi_B \cap (\Omega - C))], H(\pi_A) + H((\Omega - B) \cap \pi_C) \}$$

has to be verified,

β') if $B \cap C = \emptyset$

$$H(\pi_A) + \Psi[J(B \cup C), J(B), J(C), H(\pi_B), H(\pi_C)]$$
$$= \Psi[J(A) + J(B \cup C), J(A) + J(B), J(A) + J(C), H(\pi_A) + H(\pi_B), H(\pi_A)$$
$$+ H(\pi_C)]$$

has to be satisfied.

Setting $t = J(A)$, $x = J(B \cup C)$, $y = J(B)$, $z = J(C)$, $w = H(\pi_A)$, $u = H(\pi_B)$ $v = H(\pi_C)$, the last equality can be written in the form :

c_6') $w + \Psi(x, y, z, u, v) = \Psi(t + x, t + y, t + z, w + u, w + v)$ (consistency equation).

It cannot pass unnoticed that the consistency conditions (b_6) and (c_6) are connected with the problem of consistently extending the information measures

Next section will be devoted to the information measures which are both se. and π composite.

10. TOTALLY COMPOSITIVE INFORMATION MEASURES

Let us start with the following definition :

Definition 12. We shall say that an information measure on a given space $(\Omega,\ S,\ \varepsilon,\ \kappa^*)$ is *totally compositive* if it is set compositive and it exhibits a π composition law.

Of course the set composition law F and the π composition law Ψ of a totally compositive measure of information are not independent, in the sense that any set composition law can be coupled to any π composition law Ψ. As a matter of fact with the exception of (c_3) the properties $(c_1) - (c_6)$ are restrictions for both F and Ψ. I think that the relationship between set and π composition laws be well emphasized by the following propositions :

Proposition 7. If H is a totally compositive information measure on a given information measurable space $(\Omega,\ S,\ \varepsilon,\ \kappa^*)$ then the π composition law Ψ is reduced to a function $\Phi : \Gamma_4 \to \bar{R}^+$ with

$$\Gamma_4 = \{\ (y, z, u, v) : (F(y, z), y, z, u, v)) \in \Gamma_5\ \}$$

and F being the set composition law of H. Besides the inequality

$$\Phi(x, y, u, v) \geqslant F(x, y)$$

holds for every $(x, y, u, v) \in \Gamma_4$.

Proof : the first part of the proposition is trivially true, in fact we have $\Psi(x, y, z, u, v) = \Psi(F(y, z), y, z, u, v)) = \Phi(y, z, u, v)$; the second part is a direct consequence of (c_1).

We can observe moreover that if $\{\ A\ \} \in \varepsilon_1$, $\{\ B\ \} \in \varepsilon_1$, $A \cap B = \varnothing$ imply $\{\ A\ \} \cup \{\ B\ \} \in \varepsilon_2$, then

$$\Phi(x, y, x, y) \geqslant F(x, y),$$

for every $(x, y, x, y) \in \Gamma_4$, imply

$$\Phi(x, y, u, v) \geqslant F(x, y)$$

for every $(x, y, u, v) \in \Gamma_4$. As a matter of fact $(x, y, u, v) \in \Gamma_4$ implies in this case $(x, y, x, y) \in \Gamma_4$ and the monotonicity of H and Ψ leads to

$$u \geqslant x \quad , \quad v \geqslant y$$

hence to

$$\Phi(x, y, u, v) \geqslant \Phi(x, y, x, y).$$

We shall say that a partition π_B is a part of the partition π_A and we shall write $\pi_B \subset \pi_A$ if for every set B_i of π_B one set A_j of π_A exists such that $B_i = A_j$. It is obvious that $\{ A \} \supset \{ B \}$ implies $A = B$.

Proposition 8. If H is a totally compositive information measure on an information measurable space $(\Omega, S, \varepsilon, \kappa^*)$ such that for every

$$\pi_A \in \varepsilon : \pi_B \subset \pi_A$$

implies $\pi_B \in \varepsilon$, then the restriction of H to ε_m is a function H_m of $[J(A_1), J(A_2), ..., J(A_m)]$, $(m = 1, 2, ...)$:

$$H(\pi_A) = H_m[J(A_1), J(A_2), ..., J(A_m)]$$

for every $\pi_A = (A_1, A_2, ..., A_m) \in \varepsilon_m$ $(m = 1, 2, ...)$.

Proof : Let us denote with $\top$ the binary operation F, thus

$$F(x, y) = x \top y.$$

Then let us set

$$\overset{i=m}{\underset{i=1}{\top}} x_i = ((...(x_1 \top x_2) \top x_3) \top ...) \top x_m.$$

For every $\pi_A = (A_1, A_2, ..., A_m) \in \varepsilon_m$, we have :

$$H(\{ A_{m-1} \} \cup \{ A_m \}) = \Phi[J(A_{m-1}), J(A_m), J(A_{m-1}), J(A_m)]$$
$$= H_2[J(A_{m-1}), J(A_m)].$$

Suppose now

$$H(\{ A_2 \} \cup \{ A_3 \} \cup ... \cup \{ A_m \}) = H\left(\overset{j=m}{\underset{j=2}{\bigcup}} \{ A_j \} \right)$$

$$= H_{m-1}[J(A_2), J(A_3), ..., J(A_m)],$$

since it is

$$H(\pi_A) = H(\{ A_1 \} \cup \{ A_2 \} \cup ... \cup \{ A_m \}) = H\left[\{ A_1 \} \cup \left(\overset{j=m}{\underset{j=2}{\bigcup}} \{ A_j \} \right) \right]$$

$$= \Phi\left[J(A_1), \overset{j=m}{\underset{j=2}{\top}} J(A_j), J(A_1), H\left(\overset{j=m}{\underset{j=2}{\bigcup}} \{ A_j \} \right) \right],$$

then it is also

$$H(\pi_A) = \Phi\left[J(A_1), \mathop{\mathsf{T}}_{j=2}^{j=m} J(A_j) \quad, \quad J(A_1) \quad, \quad H_{m-1}[J(A_1), J(A_2), ..., J(A_m)] \right]$$

$$= H_m[J(A_1), J(A_2), ..., J(A_m)],$$

thus, by induction, proposition 8 is proved. However the converse is not true; this means that from

$$H(\pi_A) = H_m[J(A_1), J(A_2), ..., J(A_m)] \quad \forall \, \pi_A \in \varepsilon_m,$$

$$(m = 1, 2, ...)$$

it does not follow that H is totally compositive, even if H were set compositive. The following examples will illustrate this point :

EXAMPLE 9. Let X be a metric space. Let Ω be any non empty subset of X, S a class of non empty subsets of Ω and suppose the diameter $\delta(A) \; \forall \, A \in S$ have à minimum $\bar{\delta} > 0$ on S. As usual let ε be a class of partitions of sets $A \in S$ into a finite number of subsets $A_i \in S$ with the property : $\forall \, \pi_A \in \varepsilon$, $\pi_B \subset \pi_A \Rightarrow \pi_B \in \varepsilon$. Finally let κ^* be a collection of classes κ_{in} of $[\sigma -]$ independent partitions in ε. Consider the function $H : \varepsilon \to \bar{R}^+$ defined by

$$H(\pi_A) = \frac{1}{\bar{\delta}} \sum_{i=1}^{i=m} \left[\frac{\delta(A_i)}{\bar{\delta}} \right]^{m-2},$$

for every $\pi_A = (A_1, A_2, ..., A_m) \in \varepsilon$. The function H is an information measure on $(\Omega, S, \varepsilon, \kappa^*)$, with a suitable choice of κ^*, which could be empty. This information measure is also a function H_m of the informations $J(A_i) = \dfrac{1}{\delta(A_i)}$ of the single events A_i, but it is neither set nor π compositive.

EXAMPLE 10. Let (X, S, μ) be a measure space. Let Ω be a non empty subset of X, S a class of non empty subsets of Ω; suppose the measure $\mu(A) \; \forall \, A \in S$ have a minimum $\bar{\mu} > 0$ on S. ε and κ^* be defined as in the former example 9. Consider now the function $H : \varepsilon \to \bar{R}^+$ defined by

$$H(\pi_A) = \frac{1}{\bar{\mu}} \sum_{i=1}^{i=m} \left[\frac{\mu(A_i)}{\bar{\mu}} \right]^{m-2},$$

for every $\pi_A = (A_1, A_2, ..., A_m) \in \varepsilon$. With a suitable choice of κ^*, which could be also empty, the space $(\Omega, S, \varepsilon, \kappa^*, H)$ is an information space. But in this case the information measure H is set compositive whereas it is not π compositive and it exhibits again the property

$$H(\pi_A) = H_m[J(A_1), J(A_2), ..., J(A_m)] \; \forall \, \pi_A \in \varepsilon_m \quad (m = 1, 2, ...).$$

6

About the connections between set and π composition laws we have further the following two propositions whose proofs we shall omit :

Proposition 9. If $(\Omega, S, \varepsilon, \kappa^*)$ is an information measurable space such that for every $\pi_A \in \varepsilon : \pi_B \subset \pi_A$ implies $\pi_B \in \varepsilon$, if H is an information measure on $(\Omega, S, \varepsilon, \kappa^*)$ with the properties : 1) it is set compositive, 2) for every $\pi_A \in \varepsilon$, $H(\pi_A) = J(A)$, then H is totally compositive with π composition law

$$\Phi(x, y, u, v) = \Phi(x, y, x, y) = F(x, y),$$

where F denotes the set composition law of H.

Conversely :

Proposition 10. If $(\Omega, S, \varepsilon, \kappa^*)$ is an information measurable space such that for every $\pi_A \in \varepsilon : \pi_B \subset \pi_A$ implies $\pi_B \in \varepsilon$, if H is a totally compositive measure of information on $(\Omega, S, \varepsilon, \kappa^*)$ and

$$\Phi(x, y, x, y) = F(x, y) \quad \forall \ (x, y, x, y) \in \Gamma_4,$$

where Φ and F are the π composition law and the set composition law of H, then

$$H(\pi_A) = J(A) \qquad \forall \ \pi_A \in \varepsilon$$

holds.

11. UNIVERSAL COMPOSITION LAWS

To achieve our aim that is a new characterization for Shannon's and Rényi's measures of information by means of their composition laws we introduce now the notion of universal composition law.

Definition 13. We shall say *universal* a set composition law $F_{un}(x, y)$ if for all information measures J which are consistent with F_{un} and every choice of the information measurable space $(\Omega, S, \varepsilon, \kappa^*)$ one can assign any value in $\bar{R}^+$ to the information measures J of the independent events $\{ A_r \}$ in the same class $\kappa_{in} \in \kappa^*$.

For instance universal set composition laws are the shannonian composition law

$$F_s = - c \log (e^{-x/c} + e^{-y/c}) \qquad (c > 0)$$

and the composition law

$$F_i = \mathrm{Inf} \ (x, y).$$

On the other hand as we proved (see [3], [11], [12]) the following proposition holds.

Proposition 11. The two functions F_s and F_i are the only continuous universal set composition laws.

As it concerns the π composition laws, definition 13 suggests the following.

Definition 14. We say *universal* a π composition law $\Psi_{un}(x, y, z, u, v)$ if for all information measures H which admit Ψ_{un} as π composition law and every choice of the information measurable space $(\Omega, \, \mathcal{S}, \, \varepsilon, \, \kappa^*)$ one can assign any value in $\bar{R}^+$ to the measures of information of the independent partitions π_{A_r} in the same class $\kappa_{in} \in \kappa^*$.

For instance Shannon's π composition law

$$\Psi_S = \frac{u\,e^{-y/c} + v\,e^{-z/c}}{e^{-x/c}}$$

and Rényi's π composition law

$$\Psi_R = \frac{c}{1 - \alpha} \log \frac{e^{(1-\alpha)u/c - y/c} + e^{(1-\alpha)v/c - z/c}}{e^{-x/c}}$$

are universal.

However they are not the only universal π composition laws (see [16]).

Let us denote with $\Gamma_5^{(un)}$ the domain of a given universal composition law Ψ_{un}. From definition 14 we obtain the following properties :

1) if $(\Omega, \, \mathcal{S}, \, \varepsilon, \, \kappa^*, \, H)$ is one of the information spaces with the universal composition law Ψ_{un}, and if we define

$$\Gamma_3 = \{ (x, y, z) : x = J(A \cup B), \, y = J(A), \, z = J(B), \, \{ A \} \in \varepsilon_1,$$

$$\{ B \} \in \varepsilon_1, \, \{ A \cup B \} \in \varepsilon_1, \, A \cap B = \varnothing \},$$

then for every $(x, y, z) \in \Gamma_3$:

$$u \geqslant y \quad , \quad v \geqslant z \Rightarrow (x, y, z, u, v) \in \Gamma_5^{(un)},$$

2) $\qquad\qquad (x, y, z, u, v) \in \Gamma_5^{(un)} \Rightarrow (x, z, y, v, u) \in \Gamma_5^{(un)}.$

Besides from last definition and properties (c_1)-(c_6') of section 9 we get :

$d_1)$ $\forall \, (x, y, z, u, v) \in \Gamma_5^{(un)}$: $\Psi_{un}(x, y, z, u, v) \geqslant x,$

$d_2)$ *(symmetry)* $\forall \, (x, y, z, u, v) \in \Gamma_5^{(un)}$:

$$\Psi_{un}(x, y, z, u, v) = \Psi_{un}(x, z, y, v, u),$$

$d_3)$ *(monotonicity)* $\forall \, (x, y, z, u, v') \in \Gamma_5^{(un)}$, $v'' > v'$ implies

$$\Psi_{un}(x, y, z, u, v') \leqslant \Psi_{un}(x, y, z, u, v''),$$

$d_4)$ *(associativity)* for every $x, x', x'', x_1, x_2, x_3, u_1, u_2, u_3$ in $\bar{R}^+$ such that

$$(x', x_1, x_2, u_1, u_2) \in \Gamma_5^{(un)} \quad , \quad (x'', x_2, x_3, u_2, u_3) \in \Gamma_5^{(un)},$$

$$(x, x', x_3, x', u_3) \in \Gamma_5^{(un)} \quad , \quad (x, x_1, x'', u_1, x'') \in \Gamma_5^{(un)} :$$

$$\Psi_{un}[x, x', x_3, \Psi_{un}(x', x_1, x_2, u_1, u_2), u_3]$$

$$= \Psi_{un}[x, x_1, x'', u_1, \Psi_{un}(x'', x_2, x_3, u_2, u_3)].$$

d_5) it exists one $e(= J(\varnothing)) \in \bar{R}^+$ such that

$$\forall \, x \in \bar{R}^+, \forall \, u \geqslant x \; : \; \Psi_{un}(x, x, e, u, e) = u,$$

but since $e \geqslant x \,\forall\, x \in \bar{R}^+$, we have necessarily $e = +\infty$.

d_6) (*consistency equation*)

$$\Psi_{un}(x + t, y + t, z + t, u + w, v + w) = \Psi_{un}(x, y, z, u, v) + w$$

for every $(x, y, z, u, v) \in \Gamma_5^{(un)}$, $t \geqslant 0$, $w \geqslant t$.

Examining and comparing the two definition 13 and 14 we recognize that the only connection between a universal set composition law and a universal π composition law has to be found in their domains $\Gamma_2^{(un)}$ and $\Gamma_5^{(un)}$. In principle there is no other incompatibility between universal set and π composition laws. Thus it makes sense to look for the universal π composition laws which are consistent with a given universal set composition law $F_{un}(x, y)$. In this special case we have :

$$\Gamma_5^{(un)} = \{\, (x, y, z, u, v) : (x, y) \in \Gamma_2^{(un)}, u \geqslant y, v \geqslant z \,\},$$

and :

$$\Psi_{un} = \Psi_{un}[F_{un}(y, z), y, z, u, v] = \Phi_{un}(y, z, u, v).$$

The domain of the function Φ_{un} is now the set $\Gamma_4^{(un)}$:

$$\Gamma_4^{(un)} = \{\, (x, y, u, v) : (x, y) \in \Gamma_2^{(un)}, u \geqslant x, v \geqslant y \,\}.$$

In this domain the function Φ_{un} must exhibit the following properties :

e_1) $\forall \, (x, y, u, v) \in \Gamma_4^{(un)} \; : \; \Phi_{un}(x, y, u, v) \geqslant F_{un}(x, y),$

e_2) $\forall \, (x, y, u, v) \in \Gamma_4^{(un)} \; : \; \Phi_{un}(x, y, u, v) = \Phi_{un}(y, x, v, u),$

e_3) $\forall \, (x, y, y, v') \in \Gamma_4^{(un)} \; : \; v'' \geqslant v'$ implies

$$\Phi_{un}(x, y, u, v') \leqslant \Phi_{un}(x, y, u, v''),$$

e_4) let us denote with $\bar{\Gamma}_5^{(un)}$ the domain

$$\bar{\Gamma}^{(un)} = \{\, (x, y, z) : (x, y) \in \Gamma_2^{(un)}, (y, z) \in \Gamma_2^{(un)}, (z, x) \in \Gamma_2^{(un)} \,\}$$

then for every $(x_1, x_2, x_3) \in \bar{\Gamma}_3^{(un)}$, $u_1 \geqslant x_1, u_2 \geqslant x_2, u_3 \geqslant x_3$:

$$\Phi_{un}[F_{un}(x_1, x_2), x_3, \Phi_{un}(x_1, x_2, u_1, u_2), u_3]$$
$$= \Phi_{un}[x_1, F_{un}(x_2, x_3), u_1, \Phi_{un}(x_2, x_3, u_2, u_3)],$$

e_5) $\forall \, x \in \bar{R}^+, \forall \, u \geqslant x : \Phi_{un}(x, +\infty, u, +\infty) = u,$

e_6) for every $(x, y, u, v) \in \Gamma_4^{(un)}, \forall \, t \geqslant 0, \forall \, w \geqslant t$:

$$\Phi_{un}(x + t, y + t, u + w, v + w) = \Phi_{un}(x, y, u, v) + w.$$

We shall devote the next section to the study of the case

$$F_{un} = - c \log [e^{-x/c} + e^{-y/c}],$$

that is the case of Shannon's set composition law as universal set composition law. The problem consists in characterizing the universal π composition laws which are consistent with this universal set composition law.

12. UNIVERSAL π COMPOSITION LAWS AND THE SHANNONIAN SET COMPOSITION LAW

Let us observe first that if the universal π composition law consistent with Shannon's set composition law does not depend on x, y :

$$\Phi_{un}(x, y, u, v) = \Phi_{un}^*(u, v)$$

then it must satisfy the reduced system

f_1) $\forall (u, v) \in \Gamma_2^{(un)} : \Phi_{un}^*(u, v) \geqslant 0,$

f_2) $\forall (u, v) \in \Gamma_2^{(un)} : \Phi_{un}^*(u, v) = \Phi_{un}^*(v, u),$

f_3) $\forall (u, v') \in \Gamma_2^{(un)}, \forall v'' \geqslant v' : \Phi_{un}^*(u, v') \leqslant \Phi_{un}^*(u, v''),$

f_4) $\forall (u, v, w) \in \bar{\Gamma}_3^{(un)} :$

$$\Phi_{un}^*[\Phi_{un}^*(u, v), w] = \Phi_{un}^*[u, \Phi_{un}^*(v, w)],$$

f_5) $\forall u \in \bar{R}^+ : \Phi_{un}^*(u, + \infty) = u,$

f_6) $\forall (u, v) \in \Gamma_2^{(un)}, \forall w \geqslant 0 : \Phi_{un}^*(u + w, v + w) = \Phi_{un}^*(u, v) + w.$

These properties coincide with those we derived from the definition for an universal set composition law (see [3]). Hence if we apply the results one can find in [12], we get the following proposition :

Proposition 12. The only universal π composition laws continuous in the domain $\Gamma_2^{(un)} = \{ (u, v) : e^{-u/c} + e^{-v/c} \leqslant 1, (c > 0) \}$ and consistent with Shannon's set composition law are :

$$\Phi_{un,h}^* = - h \log [e^{-u/h} + e^{-v/h}] \qquad (0 < h < c)$$

and

$$\Phi_{un,0}^* = \text{Inf} (u, v).$$

In order to make easier the problem of finding the universal π composition laws which are consistent with Shannon's law in the general case, we observe first that from (e_6) it follows :

$$\Phi_{un}(x, y, u, v) - v = \varphi(x, y, u - v),$$

that is the function $\Phi_{un}(x, y, u, v) - v$ depends on u and v only through their difference $u - v$.

Let us now set : $U = u - v$,

$$x = - c \log X \quad , \quad y = - c \log Y$$

and

$$\bar{\varphi}(X, Y, U) = \varphi(- c \log X, - c \log Y, U).$$

Because of (e_6) we have moreover :

$$\bar{\varphi}(XT, YT, U) = \bar{\varphi}(X, Y, U) \quad \forall\, T \in\,]0, 1].$$

Then let us set :

$$T = \frac{1}{X + Y} \quad \text{and} \quad p = \frac{X}{X + Y},$$

finally let us consider the function $\theta : \bar{\mathfrak{D}} \to \bar{R}$ defined by :

$$\theta(p, U) = - \bar{\varphi}(p, 1 - p, - U)$$

with

$$\bar{\mathfrak{D}} = \{ (x, y) : 0 \leqslant x \leqslant 1, - \infty \leqslant y \leqslant + \infty \}.$$

Hence we have :

$$\Phi_{un}(x, y, u, v) = - \theta\left[\frac{e^{-x/c}}{e^{-x/c} + e^{-y/c}}, v - u\right] + v.$$

While it turns out obvious that the function $\Phi_{un}(x, y, u, v)$ verifies (e_6), in order that it can verify (e_2) (symmetry) the function $\theta(x, y)$ has to satisfy the following equation :

(I) $\theta(x, y) = y + \theta(1 - x, - y)$

for every $(x, y) \in \bar{\mathfrak{D}}$.

Moreover for (e_3) $\theta(x, y)$ has to be a monotonic non decreasing function respect to y, and if (e_5) holds then it must be $\theta(0, - \infty) = 0$. Through some easy calculations we can also recognize that the π composition law Φ_{un} verifies (e_4) (associativity) iff :

$$\text{(II)}\ y + \theta\left[u, x - y + \theta\left(\frac{u + v - 1}{u}, y\right)\right]$$

$$= x + \theta\left[v, y - x + \theta\left(\frac{v + u - 1}{v}, x\right)\right]$$

holds for every (u, v, x, y) in $\mathfrak{D}_s$ being :

$$\mathfrak{D}_s = \{ (u, v, x, y) : u \in\,]0, 1], v \in\,]0, 1], u + v \geqslant 1, x \in R, y \in R \}.$$

Thus we proved the following proposition :

Proposition 13. Let $M(\bar{\mathfrak{D}})$ denote the set of functions $\theta(x, y)$ with domain $\bar{\mathfrak{D}}$ which are non decreasing respect to y. In order that a function $\Phi_{un} : \Gamma_4^{(un)} \to \bar{R}^+$ can verify $(e_1) - (e_6)$ it must be

$$(a) \qquad \Phi_{un} = -\theta\left[\frac{e^{-x/c}}{e^{-x/c} + e^{-y/c}}, v - u\right] + v,$$

where $\theta(x, y)$ is a solution in $M(\bar{\mathfrak{D}})$ of the following system of functionnal equations

$$(I) \qquad \theta(x, y) = y + \theta(1 - x, - y) \qquad \forall\ (x, y) \in \bar{\mathfrak{D}}$$

$$(II) \quad y + \theta\left[u, x - y + \theta\left(\frac{u + v - 1}{u}, y\right)\right]$$

$$= x + \theta\left[v, y - x + \theta\left(\frac{v + u - 1}{v}, x\right)\right]$$

for every $(u, v, x, y) \in \mathfrak{D}_s$,

with the (boundary) condition $\theta(0, - \infty) = 0$.

As it concerns property (e_1) it is not difficult to prove that.

Proposition 14. In order that the function Φ_{un} defined by (a) verify (e_1) it is necessary and sufficient that the inequality

$$(III) \qquad\qquad \theta(x, 0) \leqslant - c \log (1 - x)$$

be satisfied for every $x \in [0, 1]$.

Proof : Let us observe first that $(x, y) \in \Gamma_2^{(un)}$ implies

$$[x - F_{un}(x, y), y - F_{un}(x, y)] \in \Gamma_2^{(un)}$$

and

$$F_{un}[x - F_{un}(x, y), y - F_{un}(x, y)] = 0.$$

Then $(x, y, u, v) \in \Gamma_4^{(un)}$ implies

$$[x - F_{un}(x, y), y - F_{un}(x, y), u - F_{un}(x, y), v - F_{un}(x, y)] \in \Gamma_4^{(un)}.$$

On the other hand (e_6) and (e_1) lead to

$$\Phi_{un}[x - F_{un}(x, y), y - F_{un}(x, y), u - F_{un}(x, y), v - F_{un}(x, y)]$$

$$= \Phi_{un}(x, y, u, v) - F_{un}(x, y) \geqslant 0,$$

for every $(x, y) \in \Gamma_2^{(un)}$, $u \geqslant x, v \geqslant y$. In particular for $u = x$ and $v = y$ we have :

$$\Phi_{un}[x - F_{un}(x, y), y - F_{un}(x, y), x - F_{un}(x, y), y - F_{un}(x, y)] \geqslant 0.$$

In our case where $F_{un}(x, y)$ is the shannonian set composition law

$$F_{un}(x, y) = - c \log [e^{-x/c} + e^{-y/c}],$$

setting :

$$p = \frac{e^{-x/c}}{e^{-x/c} + e^{-y/c}}$$

we obtain

$$\Phi_{un}[-c\log p, -c\log(1-p), -c\log p, -c\log(1-p)] \geqslant 0 ;$$

for (a) we get finally :

$$-\theta(p, 0) - c\log(1-p) \geqslant 0,$$

thus (III) holds.

Conversely, from (III) it follows :

$$\Phi_{un}[x - F_{un}(x, y), y - F_{un}(x, y), x - F_{un}(x, y), y - F_{un}(x, y)] \geqslant 0,$$

and because of the monotonicity of $\theta(x, y)$ with respect to y and (I) which imply the monotonicity of $\Phi_{un}(x, y, u, v)$ with respect to u and v, we have

$$\Phi_{un}[x - F_{un}(x, y), y - F_{un}(x, y), u - F_{un}(x, y), v - F_{un}(x, y)]$$

$$\geqslant \Phi_{un}[x - F_{un}(x, y), y - F_{un}(x, y), x - F_{un}(x, y), y - F_{un}(x, y)] \geqslant 0.$$

But for (a) it is :

$$\Phi_{un}[x - F_{un}(x, y), y - F_{un}(x, y), u - F_{un}(x, y), v - F_{un}(x, y)]$$

$$= \Phi_{un}(x, y, u, v) - F_{un}(x, y).$$

Hence we obtain :

$$\Phi_{un}(x, y, u, v) \geqslant F_{un}(x, y) \qquad \forall\, (x, y, u, v) \in \Gamma_4^{(un)} \,]$$

this completes the proof.

It will turn also useful to note that for an *idempotent* universal π composition law Φ_{un} necessarily

$$\theta(x, 0) = 0 \qquad \forall\, x \in [0, 1].$$

In fact if we remember that the idempotence of Φ_{un} leads to

$$\Phi_{un}(x, y, u, u) = u \qquad \forall\, (x, y) \in \Gamma_2^{(un)}, \forall\, u \geqslant \text{Sup}\,(x, y),$$

from (a) we get immediately : $\theta(x, 0) = 0 \;\forall\, x \in [0, 1]$. Conversely it is quite evident that $\theta(x, 0) = 0 \;\forall\, x \in [0, 1]$ implies the idempotence for the universal π composition law Φ_{un}. Note also that for an idempotent universal π composition law (III) follows directly from last equality $\theta(x, 0) = 0$, in the sense that in this particular case (III) is trivially verified. Thus we have :

Proposition 15. An universal π composition law Φ_{un} is idempotent and verifies $(e_1) - (e_6)$ iff (a) holds, where $\theta(x, y)$ is a solution in $M(\overline{\mathfrak{D}})$ of the following system of functional equations

(I') $\theta(x, y) = y + \theta(1 - x, - y)$ $\forall \, (x, y) \in \bar{\mathfrak{D}},$

(II') $y + \theta\left[u, x - y + \theta\left(\dfrac{u + v - 1}{u}, y\right)\right]$

$$= x + \theta\left[v, y - x + \theta\left(\dfrac{v + u - 1}{v}, x\right)\right]$$

for every $(u, v, x, y) \in \mathfrak{D}_s$,

(III') $\theta(x, 0) = 0$ $\forall \, x \in [0, 1],$

with the (boundary) condition $\theta(0, - \infty) = 0$.

The two systems of functional equations (I')-(III') and (I)-(II) are studied in [14], [13] and [15] respectively. We report here the conclusions.

Proposition 16. Shannon's π composition law

$$\Phi_{un,s} = \frac{u\,e^{-x/c} + u\,e^{-y/c}}{e^{-x/c} + e^{-y/c}}$$

and Rényi's π composition law

$$\Phi_{un,r} = \frac{c}{1 - \alpha} \log \left[\frac{e^{(1-\alpha)u/c - x/c} + e^{(1-\alpha)v/c - y/c}}{e^{-x/c} + e^{-y/c}}\right]$$

are the only universal π composition laws which exhibit the following properties : continuity, idempotence, consistency with Shannon's set composition law, strictly monotonicity with respect to v.

The universal π composition laws correspond to the solutions

$$\theta_S(x, y) = xy$$

$$\theta_R(x, y) = \frac{c}{\alpha - 1} \log [1 - x + x\,e^{(\alpha - 1)y/c}]$$

of the system (I')-(III'), respectively.

Proposition 17. Shannon's and Rényi's generalized π composition laws (see [13]) and the π composition law

$$\Phi^*_{un,h} = - h \log [e^{-u,h} + e^{-v/h}]   (0 < h < c)$$

are the only universal π composition laws consistent with Shannon's set composition law, continuous together with their derivatives up to the order three in their domains $\Gamma_4^{(un)}$.

The universal π composition law $\Phi^*_{un,h}$ corresponds to the solution

$$\theta^*_{S,h}(x, y) = h \log (1 + e^{y/h})$$

of the system (I)-(II), (see again [13])

Proposition 18. The π composition law $\Phi^*_{un,h}$ and the π composition law

$$\Phi^*_{un,0} = \text{Inf } (u, v)$$

are the only universal continuous π composition laws which do not depend on x and y and are consistent with Shannon's set composition law and the presence of the empty set, i.e. (e_5). Whereas if we esclude the empty set, that is (e_5), then the set of the universal π composition law of this kind is completed by the π composition law

$$\Phi_{un,s} = \text{Sup } (u, v).$$

These last composition laws correspond to the solutions

$$\theta_{S,0}(x, y) = \frac{|y| + y}{2}$$

$$\theta_s(x, y) = \frac{- |y| + y}{2}$$

of the system (I)-(II), respectively.

13. FINAL REMARKS AND ACKNOWLEDGMENTS

Some encouraging results have been obtained (see [9] and [17]) in applying to classical problems measures of information which do not involve any probability concept. Whence the opportunity of finding as more as possible meaningful measures of information which can differ in some qualitative features. They could suggest new and interesting applications. In this order of ideas it seems to be useful to know also all the universal π composition laws which are consistent with the universal set composition law

$$F_0(x, y) = \text{Inf } (x, y) \text{ (see [16]).}$$

Here we have just illustrated the basic common properties which have to be possessed by any information measure. We have also considered some special properties which are exhibited by the classical measures of information in order to insert the classical information theory in a possible new one.

The author wishes to thank Dr P. Benvenuti and Dr N. Pintacuda for many helpful discussions.

BIBLIOGRAPHY

[1] J. KAMPÉ de FÉRIET et B. FORTE, *Information et probabilité*, C. R. Acad. Sci. Paris, 265 (1967), 110-114.

[2] J. KAMPÉ de FÉRIET et B. FORTE, *Information et probabilité*, C. R. Acad Sci., Paris, 265 (1967), 142-146.

[3] J. KAMPÉ de FÉRIET et B. FORTE, *Information et probabilité*, C. R. Acad. Sci., Paris, 265 (1967), 350-353.

[4] B. FORTE, *On the amount of information given by an experiment*, Colloquium on Information Theory, Debrecen (1967), in printing.

[5] B. FORTE et N. PINTACUDA, *Sull'informazione associata alle esperienze incomplete*, *Ann. Mat. Pura e App., in printing*. For a summary of this paper see : Information fournie par une expérience, C. R. Acad. Sci., Paris, 266 (1968), 242-245.

[6] D. K. FADDEEV, *Zum Begriff der Entropie eines endlichen Wahrscheinlichkeitsschemas*, Arbeiten zur Informations theorie, Vol. I, VEB Deutscher Verlag der Wissensenschaften, Berlin (1957), 96-90.

[7] N. PINTACUDA, « Shannon Entropy : a more general derivation », *Statistica* 26, 2 (1966), 509-524.

[8] Z. DARÓCZY, Über eine Charakterisierung der Shannonschen Entropie, *Statistica* 27, 1 (1967), 189-205.

[9] J. L. RIGAL, N. AGGARWAL et J. C. CANONGE, *Incertitude et fonction d'imprécision liée à un questionnaire sur un espace métrique*, C. R. Acad. Sci., Paris, 263 (1966) 268-270.

[10] J. KEMPÉ de FÉRIET, *Mesure de l'information fournie par un événement*, C. R. Acad. Sci., Paris, in printing.

[11] C. BAIOCCHI, *Su un sistema di equazioni funzionali connesso alla teoria dell'informazione*, Boll. U.M.I. 23, 2 (1967), 236-246.

[12] Z. DAROCZY, *Über ein Funktionalgleichungssystem der Informations theorie*, in printing.

[13] B. FORTE et P. BENVENUTI, *Su una classe di misure regolari a traccia shannoniana*, Atti del Sem, Mat. e Fis. Univ. di Modena, in printing.

[14] B. FORTE, *On a system of functional equations in information theory*, *Aequationes Mathematicae*, in printing.

[15] B. FORTE et Z. DARÓCZY, *Sopra un sistema di equazioni funzionali nella teoria dell'informazione*, Ann. Univ. di Ferrara, in printing.

[16] P. BENVENUTI, *Sulle misure di informazione compositive a traccia compositiva non shannoniana*, manuscript.

[17] M. SCHNEIDER, *Information et mesure, application aux graphes*, Sem. sur les questionnaires, Inst. Blaise-Pascal, Paris, 23 oct. 1968.