RAIRO. INFORMATIQUE THÉORIQUE

BRUNO FORTE
CARLO SEMPI

Pruning and measures of uncertainty

RAIRO. Informatique théorique, tome 12, nº 2 (1978), p. 157-168
<http://www.numdam.org/item?id=ITA_1978__12_2_157_0>

R.A.I.R.O. Informatique théorique/Theoretical Computer Science
(vol. 12, n° 2, 1978, p. 157 à 168)

PRUNING AND MEASURES OF UNCERTAINTY (1)

by Bruno FORTE (2) and Carlo SEMPI (3)

Abstract. — *We introduce a new property, pruning, for measures of uncertainty and investigate its consequences. In particular, we derive a new characterization of the additive, subadditive and symmetric entropies without assuming expansibility. Pruning provides also a new motivation for Hartley's entropy.*

1. INTRODUCTION

It is our purpose to present a new property for measures of uncertainty-pruning. It seems to us that this property has been hitherto overlooked, unjustly so, in view of its conceptual simplicity. Pruning, as an operation, consists in the elimination of one among the possible outcomes of an experiment. It is, probably, best understood when confronted with the operation of branching, its opposite, in a sense. Branching represents a refinement of the experiment in hand; it translates the request of a more precise answer to the original question. Pruning, on the other hand, represents the focusing of the investigators's attention on a part of the possible results of the experiment. That this can be done should hardly be surprising if one bears in mind (*see* [1]) that every change in the amount of one's uncertainty about a certain state occurs because information has been gained or lost, in a process leading to that state, i. e. in a chain of experiments designed to remove the original uncertainty about that state. Thus an experiment is always performed toward a goalknowledge, complete, or, more often, partial, about a state. Pruning, as a property of an entropy, stipulates that the uncertainty about the final state does not increase if one removes one of the possible outcomes.

In the theory of questionnaires, pruning means eliminating one of the answers to a question. We should like to stress that this is not restricted to answers of zero probability (in this case one would really have expansibility rather than pruning) but can be used, and in fact must be used, when one has, for some reason, lost interest in an answer of positive probability.

(1) Research supported in part by the National Research Council of Canada under Grant A-7677.
(2) Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada.
(3) Istituto di Matematica, Universitá, Lecce, Italy.

In the theory of communication pruning may represent a situation like the following. A source broadcasts a sequence of alphabet letters, each letter having a positive probability. If one knows that a given letter does not occur in a message, then that letter can be eliminated, by pruning, thereby changing the probabilities of the remaining ones.

We shall show that pruning together with additivity, subadditivity and symmetry leads to a characterization of Hartley's entropy and that if a weaker form of pruning is defined, this together with the above mentioned properties characterizes essentially the same entropies as in [2]. In the process we wholly dispense with the property of expansibility which now appears as an artificial, albeit often useful, means of relating the entropies on probability spaces of (apparently) different cardinality.

Finally we believe that pruning will have an important role to play in the characterizations of non-symmetric entropies (*see* [3, 4]). These have a higher degree of arbitrariness than symmetric entropies. Pruning will fill the gap left by symmetry.

2. ALGEBRAIC PRELIMINARIES

Let X be a non-empty set and let $\mathscr{A}$ be an algebra of subsets of X. We shall say that a set $A \subset X$ is an atom of the algebra $\mathscr{A}$ if :

(i) $A \in \mathscr{A}$;

(ii) $B \in \mathscr{A}$, $B \subset A$ imply either $B = \emptyset$ or $B = A$.

Condition (ii) says that, in an algebra $\mathscr{A}$, the empty set is the only proper subset of an atom. As a consequence any two distinct atoms A and A' of an algebra are disjoint and therefore every union of atoms is a disjoint union.

If $A_1, A_2, \ldots, A_n$ are all the atoms of an algebra $\mathscr{A}$ of subsets of X and if $X = \bigcup_{i=1}^{n} A_i$, then every non-empty set $A \in \mathscr{A}$ can be expressed as a finite union of atoms, namely $A = \bigcup_{i=1}^{n} A_{k_i}$ where $\{ k_1, k_2, \ldots, k_r \}$ is a subset of $\{ 1, 2, \ldots, n \}$. Indeed, either A is itself an atom and must therefore coincide with A_i for some index i, or it must properly contain an atom, say A_{k_1}. Applying the same argument to $B_1 = A - A_{k_1}$, one sees that either B_1 is an atom, say A_{k_2}, or B_1 properly contains an atom. In the former case $A = A_{k_1} \cup A_{k_2}$, whilst in the latter one the same argument need be applied. This procedure ends after at most n steps.

If the conditions stated above are fulfilled, that is if $X = \bigcup_{i=1}^{n} A_i$, when A_i $(i = 1, 2, \ldots, n)$ are all the atoms of $\mathscr{A}$ we shall call $\mathscr{A}$ a *finite atomic algebra* and shall say that it is *generated* by the family of atoms $\{ A_1, A_2, \ldots, A_n \}$. In the following we shall consider only finite atomic algebrae. Such algebrae are, obviously, completely determined by the set of atoms.

Example 1: Let $X = \{ 1, 2, 3, 4, 5, 6 \}$ and let $\mathscr{A} = \mathscr{P}(X)$ be the family of all the subsets of X. Then $\mathscr{A}$ is an algebra since X is finite; moreover it is a finite atomic algebra generated by the atoms $\{ 1 \}$, $\{ 2 \}$, $\{ 3 \}$, $\{ 4 \}$, $\{ 5 \}$, $\{ 6 \}$.

We shall now introduce the operations of branching and pruning.

Branching: Branching is the operation $\mathfrak{b}$ defined on the pair $(X, \mathscr{A})$ and consisting in replacing the algebra $\mathscr{A}$ by the algebra $\mathscr{A}_k^1$ generated by the atoms $A_1, A_2, \ldots, A_{k-1}, A'_k, A''_k, A_{k+1}, \ldots, A_m$ for some $k \leqq m$, where $A'_k \cup A''_k = A_k$ and neither A'_k nor A''_k belongs to $\mathscr{A}$. Then $\mathfrak{b}_k(X, \mathscr{A}) \to (X, \mathscr{A}_k^1)$ or, for short, $\mathfrak{b}_k \mathscr{A} = \mathscr{A}_k^1$; we shall say that $\mathfrak{b}_k$ represents branching on the atom A_k.

It is obvious from the above construction that $\mathscr{A} \subset \mathscr{A}_k^1$; then the operation of branching introduces a relationship of partial ordering in the class of finite atomic algebrae of subsets of a set X.

Pruning: Pruning is defined as the operation $\mathfrak{p}$ consisting in suppressing an atom of the algebra $\mathscr{A}$. By means of pruning the pair $(X, \mathscr{A})$ is replaced by the pair $(A, \mathscr{A}_A)$ when $A = \bigcup_{i \neq k} A_i$, if A_k is the atom that has been eliminated, and $\mathscr{A}_A = A \cap \mathscr{A}$ is a subalgebra of $\mathscr{A}$. Formally we shall write $\mathfrak{p}_k : (X, \mathscr{A}) \to (A, \mathscr{A}_k)$ and shall say that $\mathfrak{p}_k$ represents the pruning of the k-th atom.

The application of the pruning operation can be repeated so as to replace $(X, \mathscr{A})$ by $(B, \mathscr{A}_B)$ where B is the union of a finite number $j < m$ of atoms of $\mathscr{A}$ and $\mathscr{A}_B = B \cap \mathscr{A}$.

Pruning induces a relationship of partial ordering on the family of finite atomic algebrae of subsets $A \in \mathscr{A}$. We shall write, in the notation just introduced

$$(A, \mathscr{A}_A) = \mathfrak{p}_k(X, \mathscr{A}) < (X, \mathscr{A}).$$

3. UNCERTAINTY, BRANCHING AND PRUNING

Let J be a compositive measure of information defined on $(X, \mathscr{A})$ (*see* [5]) and let H be a measure of expected information consistent with J (*see* [6]). As H depends on the pair $(X, \mathscr{A})$ we shall write $H = H(X, \mathscr{A})$.

Because every set of $\mathscr{A}$ is a union of atoms and because J is compositive it is actually possible to write $H = H_n(A_1, A_2, \ldots, A_n)$ if $\{A_1, A_2, \ldots, A_n\}$ is the set of atoms that generates $\mathscr{A}$. H measures the uncertainty associated with the experiment that intends to determine the $\mathscr{A}$-atoms of X.

In connection with branching we should expect that a refinement of the experiment, such as represented by branching will increase or, at least, will not decrease the uncertainty about the outcomes; this leads to the following requirement

$$H(X, \mathscr{A}) \leqq H[\mathfrak{b}_k(X, \mathscr{A})]. \tag{3.1}$$

By (3.1), uncertainty is non-decreasing with respect to the order relationship induced by branching.

As for pruning, if preliminary runs of the experiment were to lead to restrict one's attention to a subset A of X, this would come about because new information regarding the outcomes of the experiment has been received. One should therefore expect that the uncertainty one initially had will not have increased. It is thus natural to require that the uncertainty H be a non-decreasing function with respect to the order relation induced by pruning; more precisely we shall require that

$$H(X, \mathscr{A}) \geqq H[\mathfrak{p}(X, \mathscr{A})]. \tag{3.2}$$

Example 2: Let X and $\mathscr{A}$ be as in example 1 and let $\mathscr{A}'$ be generated by the atoms $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{5, 6\}$. Applying the branching operation to the atom $\{5, 6\}$ of $\mathscr{A}'$ and writing $\{5, 6\} = \{5\} \cup \{6\}$ one obtains $\mathfrak{b}\,\mathscr{A}' = \mathscr{A}$.

Example 3: Let X and $\mathscr{A}$ be as in example 1. Let

$$A = \{\text{the even numbers in } X\} = \{2, 4, 6\}$$

and let $\mathfrak{p}_i$ represent the pruning of the i-th atom $\{i\}$. Then

$$A = \{2\} \cup \{4\} \cup \{6\} \qquad \text{and} \qquad \mathfrak{p}_3\,\mathfrak{p}_2\,\mathfrak{p}_1(X, \mathscr{A}) = (A, \mathscr{A}_A),$$

where $\mathscr{A}_A$ is the algebra generated by $\{2\}$, $\{4\}$, $\{6\}$.

4. BRANCHING, ADDITIVITY, SUBADDITIVITY

Let $\mathscr{A}$ and $\mathscr{B}$ be the finite atomic algebrae of X generated by the families of atoms $\{A_1, A_2, \ldots, A_n\}$ and $\{B_1, B_2, \ldots, B_m\}$ respectively. We shall call *product atomic algebra* of $\mathscr{A}$ and $\mathscr{B}$, and shall denote it by $\mathscr{A} \times \mathscr{B}$, the algebra of subsets of X generated by the non-empty intersections $C_{ij} = A_i \cap B_j$ ($C_{ij} \neq \varnothing$).

Example 4: Let X be the same set as in example 1, let $\mathscr{A}''$ be the algebra generated by the atoms $\{1, 2\}$, $\{3, 4\}$, $\{5, 6\}$ and let $\mathscr{B}$ be generated by $\{1, 3, 5\}$ and $\{2, 4, 6\}$. Then $\mathscr{A}'' \times \mathscr{B}$ is generated by the atoms $\{i\}$ ($i = 1, 2, \ldots, 6$), so that the algebra $\mathscr{A}$ of example 1 is the product atomic algebra $\mathscr{A}'' \times \mathscr{B}$. Notice that $\mathscr{A}''$ and $\mathscr{B}$ are algebraically independent (*see* [7]), that is $C_{ij} = A_i \cap B_j \neq \emptyset$ $(1 \leqq i \leqq n, 1 \leqq j \leqq m)$.

Clearly one has both $\mathscr{A} \subset \mathscr{A} \times \mathscr{B}$ and $\mathscr{B} \subset \mathscr{A} \times \mathscr{B}$. Moreover the product algebra $\mathscr{A} \times \mathscr{B}$ can be thought of as obtained from either $\mathscr{A}$ or $\mathscr{B}$ by repeated application of branching. This is most easily verified when $m = 2$, although it can be proved in the same way for an arbitrary positive integer m. Starting from $\mathscr{A}$ one has

$$\begin{aligned}
\mathfrak{b}_1 \mathscr{A} &= \{A_1 \cap B_1, A_1 \cap B_2, A_2, \ldots, A_n\} = \mathscr{A}_1, \\
\mathfrak{b}_2 \mathscr{A}_1 &= \{A_1 \cap B_1, A_1 \cap B_2, A_2 \cap B_1, A_2 \cap B_2, A_3, \ldots, A_n\} = \mathscr{A}_2, \\
&\vdots \\
\mathfrak{b}_n \mathscr{A}_{n-1} &= \{A_1 \cap B_1, A_1 \cap B_2, \ldots, A_{n-1} \cap B_2, A_n \cap B_1, A_n \cap B_2\} \\
&= \mathscr{A} \times \mathscr{B},
\end{aligned}$$

where, at each step, the algebra obtained by branching has been identified with the set of atoms that generates it and where the empty intersections $A_i \cap B_j = \emptyset$, if any, have been discarded. Then, because of (3. 1), one has both

$$H(X, \mathscr{A}) \leqq \mathscr{H}(X, \mathscr{A} \times \mathscr{B}) \tag{4.1}$$

and

$$H(X, \mathscr{B}) \leqq \mathscr{H}(X, \mathscr{A} \times \mathscr{B}). \tag{4.2}$$

Furthermore it would seem natural to require that the uncertainty associated with the experiment $(X, \mathscr{A} \times \mathscr{B})$ be not greater than the sum of the uncertainties associated with $(X, \mathscr{A})$ and $(X, \mathscr{B})$ and that it should be equal to that sum if the algebrae $\mathscr{A}$ and $\mathscr{B}$ are algebraically independent. We should therefore expect the uncertainty H to satisfy, beside (4.1) and (4.2) the two further properties :

$$\left.\begin{array}{c} (subadditivity): \\ \mathscr{H}(X, \mathscr{A} \times \mathscr{B}) \leqq H(X, \mathscr{A}) + H(X, \mathscr{B}) \end{array}\right\} \tag{4.3}$$

and

$$\left.\begin{array}{c} (additivity): \\ \mathscr{A}, \mathscr{B} \text{ independent} \Rightarrow \mathscr{H}(X, \mathscr{A} \times \mathscr{B}) = H(X, \mathscr{A}) + H(X, \mathscr{B}). \end{array}\right\} \tag{4.4}$$

Neither (4.3) nor (4.4) is a consequence of (4.1) and (4.2). If $H(X, \mathscr{B}) > -\infty$ for at least one finite atomic algebra $\mathscr{B}$ of subsets of X, then (4.2) and (4.3) entail $H(X, \mathscr{A}) \geqq 0$ for every finite atomic algebra $\mathscr{A}$ of subsets of X.

In order to render the assumptions (4.3) and (4.4) plausible, it suffices to consider the special, but important, case in which H depends only on the cardinality of set of atoms of $\mathscr{A}$: $H(X, \mathscr{A}) = h(n)$. Then if $\mathscr{A}$ and $\mathscr{B}$ have cardinalities respectively equal to n and to m, the cardinality of their product atomic algebra $\mathscr{A} \times \mathscr{B}$ is at most equal to mn and equals mn if $\mathscr{A}$ and $\mathscr{B}$ are algebraically independent.

5. PROBABILITY ALGEBRAE

We shall now confine our considerations to measures of uncertainty, or entropies, that depend only on probabilities.

A probability measure P on $(X, \mathscr{A})$ is completely determined if the values $p_i = P(A_i)\,(i = 1, 2, \ldots, n)$ that P takes on the atoms $A_1, A_2, \ldots, A_n$ of $\mathscr{A}$ are known; for if $A \in \mathscr{A}$, $A \neq \emptyset$, then A can be expressed as a disjoint union of atoms of the algebra $\mathscr{A}$, $A = \bigcup_{j=1}^{r} A_{k_j}$ and hence $P(A) = \sum_{j=1}^{r} p_{k_j}$.

Thus the set

$$\Gamma'_n := \left\{ (p_1, p_2, \ldots, p_n) : p_i > 0\,(i = 1, 2, \ldots, n),\ \sum_{i=1}^{n} p_i = 1 \right\}$$

corresponds to the totality of probability measures P defined on $(X, \mathscr{A})$ through $p_i = P(A_i)\,(i = 1, 2, \ldots, n)$. Notice that by considering Γ'_n rather than its closure

$$\Gamma_n := \left\{ (p_1, p_2) \ldots, p_n) : p_i \geqq 0\,(i = 1, 2, \ldots, n),\ \sum_{i=1}^{n} p_i = 1 \right\}$$

the possibility that an atom of $\mathscr{A}$ may have probability equal to zero is explicitly ruled out.

It is now an easy task to translate the properties of the uncertainty into the language and the notation of probability.

If a branching operation is performed on the atom A_k so that $A_k = A'_k \cup A''_k$, $p'_k = p(A'_k)$ and $p''_k = P(A''_k)$, then the branching inequality (3.1) yields

$$\begin{aligned} &H_n(p_1, p_2, \ldots, p_{k-1}, p_k, p_{k+1}, \ldots, p_n) \\ &\quad \leqq H_{n+1}(p_1, p_2, \ldots, p_{k-1}, p'_k, p''_k, p_{k+1}, \ldots, p_n), \end{aligned}$$

for every $(p_1, p_2, \ldots, p_{k-1}, p'_k, p''_k, p_{k+1}, \ldots, p_n) \in \Gamma'_{n+1}\ (n = 2, 3, \ldots)$, with $p'_k + p''_k = p_k$. This form of branching is weaker than the usual one where one assumes the difference (*see* [5]) :

$$H_{n+1}(p_1, p_2, \ldots, p_{k-1}, p'_k, p''_k, p_{k+1}, \ldots, p_n)$$
$$-H_n(p_1, p_2, \ldots, p_{k-1}, p_k, p_{k+1}, \ldots, p_n),$$

to be a function of p'_k, p''_k and, possibly, n.

Subadditivity takes the usual form

$$\begin{aligned} &H_{mn}(p_{11}, p_{12}, \ldots, p_{1n}, p_{21}, p_{22}, \ldots, p_{2n}, \ldots, p_{m1}, p_{m2}, \ldots, p_{mn}) \\ &\leqq H_m\left(\sum_{j=1}^{n} p_{ij}, \sum_{j=1}^{n} p_{2j}, \ldots, \sum_{j=1}^{n} p_{mj}\right) \\ &+H_n\left(\sum_{i=1}^{m} p_{i1}, \sum_{i=1}^{m} p_{i2}, \ldots, \sum_{i=1}^{m} p_{in}\right), \end{aligned} \tag{5.1}$$

for every $(p_{11}, p_{12}, \ldots, p_{mn}) \in \Gamma'_{mn}$ $(m, n = 2, 3, \ldots)$.

Since two stochastically independent algebrae are also algebraically independent, the additivity property as well has the usual form

$$\begin{aligned} &H_{mn}(p_1, q_2, p_1 q_2, \ldots, p_1 q_n, p_2 q_1, \\ &\qquad p_2 q_2, \ldots, p_2 q_n, \ldots, p_m q_1, p_m q_2, \ldots, p_m q_n) \\ &= H_m(p_1, p_2, \ldots, p_m)+H_n(q_1, q_2, \ldots, q_n), \end{aligned} \tag{5.2}$$

for every $(p_1, p_2, \ldots, p_m) \in \Gamma'_m$ and every $(q_1, q_2, \ldots, q_n) \in \Gamma'_n$ $(m, n = 2, 3, \ldots)$.

If a pruning operation is performed on the set A_{n+1} and if $A = \bigcup_{i=1}^{n} A_i$, the probabilities of the atoms of $\mathscr{A}_A$ must be replaced by their respective conditional probabilities given A. That is one must replace $p_i = P(A_i)$ by $q_i = P(A_i/A) = p_i/p$ $(i = 1, 2, \ldots, n)$, where $p = P(A) = \sum_{i=1}^{n} p_i$ in order that the vector $(q_1, q_2, \ldots, q_n)$ may belong to Γ'_n. The pruning inequality (3.2) reads now

$$H_n(q_1, q_2, \ldots, q_n) = H_n\left(\frac{p_1}{p}, \frac{p_2}{p}, \ldots, \frac{p_n}{p}\right) \leqq H_{n+1}(p_1, p_2, \ldots, p_{n+1}), \tag{5.3}$$

for every $(p_1, p_2, \ldots, p_{n+1}) \in \Gamma'_{n+1}$ $(n = 2, 3, \ldots)$.

One of the properties usually postulated for the uncertainty is symmetry. Indeed, it seems to be perfectly natural to ask that

$$H_n(p_1, p_2, \ldots, p_n) = H_n(p_{\pi(1)}, p_{\pi(2)}, \ldots, p_{\pi(n)}), \tag{5.4}$$

for every $(p_1, p_2, \ldots, p_n) \in \Gamma'_n$ $(n = 2, 3, \ldots)$, where $(\pi(1), \pi(2), \ldots, \pi(n))$ is any permutation of $(1, 2, \ldots, n)$. As a permutation of the indeces amounts to a relabelling of the atoms, it should not affect $H(X, \mathscr{A})$; therefore we shall henceforth assume symmetry (5.4).

A property that we shall not use in the following, but one that we shall mention is expansibility:

$$H_{n+1}(p_1, p_2, \ldots, p_n, 0) = H_n(p_1, p_2, \ldots, p_n), \tag{5.5}$$

for every $(p_1, p_2, \ldots, p_n) \in \Gamma_n$ $(n = 2, 3, \ldots)$.

6. CONSEQUENCES OF PRUNING

A sequence of entropies $H_n : \Gamma'_n \to R$ $(n = 2, 3, \ldots)$ exhibits the properties of subadditivity (5.1), additivity (5.2) and symmetry (5.4) if, and only if (*see* [2]):

$$H_n(p_1, p_2, \ldots, p_n) = a\, H_n^S(p_1, p_2, \ldots, p_n) + A(n), \tag{6.1}$$

for every $(p_1, p_2, \ldots, p_n) \in \Gamma'_n$ $(n = 2, 3, \ldots)$, where $a \geqq 0$, H_n^S represents Shannon's entropy

$$H_n^S(p_1, p_2, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

and where A is a completely additive number theoretical function, i. e. $A(mn) = A(m) + A(n)$ $(m, n = 2, 3, \ldots)$.

The further property of expansibility (5.5) postulated in [2] was required to prove that $A(n) = b \log_2 n$ with $b \geq 0$ and then to extend the representation obtained to Γ_n.

If $p \in \,]0, 1[$ and $(q_1, q_2, \ldots, q_n) \in \Gamma'_n$ let $A_1, A_2, \ldots, A_{n+1}$ be all the atoms of an algebra $\mathscr{A}$ of subsets of X. The probabilities of the atoms A_i are taken to be

$$P(A_i) = pq_i \quad (i = 1, 2, \ldots, n), \qquad P(A_{n+1}) = 1 - p.$$

Let $A = \bigcup_{i=1}^{n} A_i$ and let us consider the pruning operation $\mathfrak{p}(X, \mathscr{A}) = (A, \mathscr{A}_A)$.

Since $P(A) = \sum_{i=1}^{n} P(A_i) = p \sum_{i=1}^{n} q_i = p$ one has $P(A_i/A) = q_i$ $(i = 1, 2, \ldots, n)$ and therefore from (5.3):

$$H_n(q_1, q_2, \ldots, q_n) \leqq H_{n+1}(pq_1, pq_2, \ldots, pq_n, 1-p), \tag{6.2}$$

or, using (6.1):

$$\begin{aligned} &a\, H_n^S(q_1, q_2, \ldots, q_n) + A(n) \\ &\quad \leqq -a \sum_{i=1}^{n} pq_i \log pq_i - a(1-p)\log(1-p) + A(n+1) \\ &\quad = a\, H_2^S(p, 1-p) + ap\, H_n^S(q_1, q_2, \ldots, q_n) + A(n+1). \end{aligned}$$

The last inequality can be written in the form

$$a H_2^S(p, 1-p) - a(1-p) H_n^S(q_1, q_2, \ldots, q_n) + A(n+1) - A(n) \geqq 0, \quad (6.3)$$

for every $p \in]0, 1[$ and for every $(q_1, q_2, \ldots, q_n) \in \Gamma'_n$. The l. h. s. of (6.3) is a continuous function of p on the closed interval $[0, 1]$; and since the maximum, $\log n$, of $H_n^S(q_1, q_2, \ldots, q_n)$ is attained for $q_i = 1/n$ $(i-1, 2, \ldots, n)$ it follows from (6.3) that

$$A(n+1) - A(n) \geqq a \log n \geqq 0, \quad (6.4)$$

must hold true. This inequality shows that pruning is compatible with subadditivity, additivity and symmetry only in the case $a = 0$. In fact, if $a > 0$, it follows from (6.4) that $\liminf_{n \to \infty} [A(n+1) - A(n)] = +\infty$; in particular, then, $\liminf_{n \to \infty} [A(n+1) - A(n)] \geqq 0$, and Kátai's theorem [9] ensures that $A(n) = b \log n$ with $b \geqq 0$ $(n = 2, 3, \ldots)$. But then (6.4) is impossible if $a > 0$. Thus $a = 0$. In this case it follows from (6.4) that

$$A(n+1) - A(n) \geqq 0 (n-2, 3, \ldots).$$

A well-known theorem [10] gives $A(n) = b \log n$ with $b \geqq 0$, so that the representation (6.1) reduces, up to the constant b, to Hartley's entropy

$$H_n^H(p_1, p_2, \ldots, p_n) = \log_2 n,$$

for every $(p_1, p_2, \ldots, p_n) \in \Gamma'_n$ $(n = 2, 3, \ldots)$. We have thus established the following characterization of Hartley's entropy.

THEOREM 1: *If the sequence $H_n : \Gamma'_n \to R$ exhibits the properties of subadditivity* (5.1) *additivity* (5.2), *pruning* (5.3), *symmetry* (5.4) *and normalization* $H_2(1/2, 1/2) = 1$, *then, and only then*

$$H_n(p_1, p_2, \ldots, p_n) = \log_2 n,$$

for every $(p_1, p_2, \ldots, p_n) \in \Gamma'_n$.

It is important to stress that no use of expansibility (5.5) has been made in proving theorem 1.

Since we do not want to forsake an entropy as rich in applications as Shannon's and as, at the same time, we should like to include pruning among the natural properties of entropy, we shall assume a weaker form of pruning than (5.3).

We shall say that a sequence of entropies $H_n : \Gamma'_n \to R$ exhibits the property of weak pruning, or, for short, that it is weakly pruning, if for each $n \geqq 2$

a real number $\bar{p}_n \in]0, 1[$ exists such that the inequality (6.2) is satisfied for every $(q_1, q_2, \ldots, q_n) \in \Gamma'_n$ and for every $p \in [\bar{p}_n, 1[$.

It is not difficult to verify that Shannon's entropy is weakly pruning. In fact, (6.2) reads

$$H_2^S(p, 1-p)-(1-p)H_n^S(q_1, q_2, \ldots, q_n) \geqq 0.$$

The minimum of the l. h. s. as $(q_1, q_2, \ldots, q_n)$ varies in Γ'_n is

$$H_2^S(p, 1-p)-(1-p)\log n.$$

Now it suffices to set $\bar{p}_n$ equal to the unique solution of the equation

$$H_2^S(p, 1-p)-(1-p)\log n = 0,$$

in the open interval $]0, 1[$, in order to have $H_2^S(p, 1-p)-(1-p)\log n > 0$ for $p \in]\bar{p}_n, 1[$.

It will be shortly shown that weak pruning is compatible with the representation (6.1).

Let

$$\varphi(n) := \inf_{p \in]0, 1[} \{p : H_{n+1}(pq_1, pq_2, \ldots, pq_n, 1-p) \geqq H_n(q_1, q_2, \ldots, q_n), (q_1, q_2, \ldots, q_n) \in \Gamma'_n\}. \tag{6.5}$$

It follows from (6.5) and (6.3), which is direct consequence of (6.2), that

$$a H_2^S[\varphi(n), 1-\varphi(n)] - a[1-\varphi(n)] H_n^S(q_1, q_2, \ldots, q_n)+A(n+1)-A(n) \geqq 0, \tag{6.6}$$

for every $(q_1, q_2, \ldots, q_n) \in \Gamma'_n$.

Our next result concerns the asymptotic behavior of φ.

THEOREM 2: *The limit of* $\varphi(n)$ *as* n *tends to infinity exists and either* $\lim_{n\to\infty} \varphi(n) = 0$ *or* $\lim_{n\to\infty} \varphi(n) = 1$.

Proof: It was shown above that if $a = 0$ then $\varphi(n) = 0$ for $(n = 2, 3, \ldots)$; in this case $\lim_{n\to\infty} \varphi(n) = 0$.

Let us consider now the case $a > 0$ and set $q_i = 1/n\ (i = 1, 2, \ldots, n)$. Then (6.6) yields

$$A(n+1)-A(n) \geqq a[1-\varphi(n)]\log n - a H_2^S[\varphi(n), 1-\varphi(n)]. \tag{6.7}$$

Define $p' := \liminf_{n\to\infty} \varphi(n)$. If $p' < 1$, it follows from (6.7) that

$$\liminf_{n\to\infty}[A(n+1)-A(n)] = +\infty.$$

The same argument as was used above gives $A(n) = b \log n$ with $b \geqq 0$; but even with this representation of $A(n)$, (6.7) leads to a contraddiction if $p' < 1$. Therefore $p' \geqq 1$ and since $\limsup_{n \to \infty} \varphi(n) \leqq 1$, the theorem is proven.

It follows from the proof of the theorem that the convergence of φ to 1 must be such that $\lim_{n \to \infty} [1 - \varphi(n)] \log n$ is a (non-negative) constant. But then again the inequality (6.7) yields $\liminf_{n \to \infty} [A(n+1) - A(n)] \geqq 0$ and therefore $A(n) = b \log n$ with $b \geqq 0$. Thus the representation (6.1) becomes

$$H_n(p_1, p_2, \ldots, p_n) = a H_n^S(p_1, p_2, \ldots, p_n) + b \log n \qquad (6.8)$$

with $a \geqq 0$, $b \geqq 0$ for every $(p_1, p_2, \ldots, p_n) \in \Gamma'_n$ $(n = 2, 3, \ldots)$.

THEOREM 3: *If $\varphi(n)$ $(n = 2, 3, \ldots)$ is defined by* (6.5) *the inequality*

$$a H_n^S(p, 1-p) - a(1-p) H_n^S(q_1, q_2, \ldots, q_n) + b \log(n+1) - b \log n \geqq 0, \qquad (6.9)$$

with $a \geqq 0$, $b \geqq 0$, is satisfied by every $(q_1, q_2, \ldots, q_n) \in \Gamma'_n$ and by every $p \in]\varphi(n), 1]$.

Proof: It suffices to deal with the case $a > 0$, $b > 0$ since it has already been established that the result is true if either a or b (or both) equals zero. The minimum of the l. h. s. of (6.9) as $(q_1, q_2, \ldots, q_n)$ varies in Γ'_n is attained for $q_i = 1/n$ $(i = 1, 2, \ldots, n)$; so we shall consider the inequality

$$a H_2^S(p, 1-p) - a(1-p) \log n + b \log(n+1) - b \log n \geqq 0. \qquad (6.10)$$

Assume $\varphi(n) \geqq n/(n+1)$; the inequality (6.10) is certainly satisfied for $p = 1$ and since the l. h. s. of (6.10), as a function of p, does not increase in the interval $[n/(n+1), 1]$, it is satisfied for $p \in]\varphi(n), 1]$. If $\varphi(n) < n/(n+1) < p$ the same argument applies. If $\varphi(n) < p < n/(n+1)$, the assertion follows from the observation that, as a function of p, the l. h. s. of (6.10) is non-decreasing in $[0, n/(n+1)]$.

COROLLARY: *The entropies* (6.8) *are weakly pruning.*

Proof: It suffices to take $\bar{p}_n > \varphi(n)$. The result then follows from theorem 3.

The results of this section can now be collected in the following characterization theorem;

THEOREM 4: *If the entropies $H_n : \Gamma'_n \to R$ are subadditive* (5.1), *additive* (5.2) *symmetric* (5.4) *and weakly pruning, then, and only then, H_n has the representation* (6.8).

REFERENCES

1. H. CICILEO and B. FORTE, *Measures of Ignorance, Information and Uncertainty*, Calcolo, Vol. 8, 1976, pp. 215-236.
2. J. ACZÉL, B. FORTE and C. T. NG, *Why the Shannon and Hartley Entropies are « Natural »*, Advances in Appl. Probability, Vol. 6, 1974, pp. 131-146.
3. B. FORTE, *Subadditive Entropies for a Random Variable*, Boll. Un. Mat. It. Vol. 14-B, 1977, pp. 118-133.
4. B. FORTE and C. A. BORTONE, *Non-symmetric Entropies with the Branching Property*, Utilitas Math. Vol. 12, 1977, pp. 3-23.
5. J. KAMPÉ DE FÉRIET et B. FORTE, *Information et Probabilité*, C.R. Acad. Sc. Paris, Série A, t. 265, 1967, pp. 110-116.
6. B. FORTE, *Measures of Information : the General Axiomatic Theory*, R.A.I.R.O., série R-3, nº 2, 1969, pp. 63-90.
7. D. A. KAPPOS, *Strukturtheorie der Wahrscheinlichkeitsfelder und räume*, Springer Verlag, Berlin, 1960.
8. J. ACZÉL and Z. DARÓCZY, *On Measures of Information and Their Characterizations*, Academic Press, New York, 1975.
9. T. KÁTAI, *A Remark on Additive Arithmetical Functions*, Ann. Univ. Sci., Budapest, Eötvös Sect. Math., Vol. 10, 1967, pp. 81-83.
10. P. ERDÖS, *On the Distribution Function of Additive Functions*, Ann. of Math., Vol. 47, 1946, pp. 1-20.