IRE Transactions on INFORMATION THEORY

Vol. IT-2, No. 3 September, 1956

1956 SYMPOSIUM ON INFORMATION THEORY

held at

Massachusetts Institute of Technology
Cambridge, Massachusetts

September 10-12, 1956

PUBLISHED BY THE
Professional Group on Information Theory

ON an APPLICATION of SEMI GROUPS METHODS TO SOME PROBLEMS in CODING

By M.P. Schützenberger
(C.N.R.S. Paris)

0. Introduction.

The current paper deals with a chapter in what could be called communication theory in extensive form : it starts with extremely restricted structures and it stops where begins the canonical problem of optimalisation. It even ends sooner for no full use of the definitions is made and the main ergodic theorem is stated without proof.

Actually the nature itself of the question under study has commanded these restrictions together with the architecture of the paper : we give a abstract model of some sort of language and we try to show how semi group concepts apply fruitfully to it with the hope that some of them may be at least of stimulating interest to specialists working on natural languages.

As frequent in the field of cybernetic, the mathematics involved even if quite simple are far away from classical analysis and, indeed, many of the necessary tools had to be sharpened especially for the purpose.

Thus the paper is twofold : in a first part the model and its main properties are discussed at a concrete level on the simplest cases : the coding and decoding with length bounded codes. In a second part a selection of theorems are proved whenever the necessary semi group theoretic preliminaries are not exacting. The link along this tail of appendices is the theory developped verbally in the first part. Finally a special chapter provides a bridge toward probalistic applications.

It is proper at this place to acknowledge the contributions of three authors who influenced deeply the building of the theory :

Sardinas and Patterson(1) who discussed first on a logical basis the general coding process.

B. Mandelbrot(2) who recognised and studied extensively the role of "word units" in communication theory and related the problem to Feller's recurrent events.

P. Dubreuil(3) and his school whose pionnering work on discrete semi groups has provided many basic concepts and arguments as it will be seen below.

Part I

1. Preliminary definition of a discrete semi group language :

We shall be concerned with the two basic sets of communication theory :

The set of all messages which may possibly be sent.

The set of all signals available for transmission along the line.

The main feature of the theory is the postulationnal requirement that the signals as well as the messages pertain both to some common class of structures so that coding and decoding not only be inverse operations but far more generally, be special instances of a quite broad new process, that of translation.

This identity of structure itself between two sets is a result from the basic restriction that they develop homogeneously in time - or more accurately that both admit a common partial order and composition operation.

That such requirements are rather stringent is clearly seen by the exemple of photography (two exposures give rarely a result which is, in any sense, equivalent to a third one) or even by harmonic modulation where Fourier transform exchanges so well time and frequency that finite signals cannot be fully adequate.

On the other hand, languages either spoken, written or gesticulated are somewhat akin with our consideration, and we shall use the name of "discrete semi group languages" (d.s.g.l.) for naming the elemental concepts of our study.

The definitions below are quite general and as said before, no full use of them will be made here - very little gain in simplicity would be achieved by using more restrictive ones.

DEFINITIONS :

I. A discrete semi group language will be a set Λ of object called "messages" satisfying the following conditions :

I.1. If λ_i and λ_j pertain to Λ so does their "product" $\lambda_k = \lambda_i \lambda_j$ made up of "λ_i" followed by "λ_j" (λ_i will be said a left divisor and λ_j a right divisor of λ_k).

I.2. If λ_i , λ_j and λ_k pertain to Λ and if $\lambda_l = \lambda_i \lambda_j$ and $\lambda_m = \lambda_j \lambda_k$ then $\lambda_i \lambda_m$ is identical with $\lambda_l \lambda_k$.

I.3. The "vacuous message" φ pertains to Λ and satisfies $\varphi \lambda_i = \lambda_i \varphi = \lambda_i$ for all $\lambda_i \in \Lambda$

I.4. There is a sub set Λ_c from Λ called "dictionary" or "basis" whose elements are called "words". Λ_c is such as :

I.4.1. φ does not pertain to Λ_c

I.4.2 for all $\lambda_i \in \Lambda - \varphi$

either $\lambda_i \in \Lambda_c$

either these exist a unique finite set of words $\lambda_{i_1}, \lambda_{i_2}, \dots \lambda_{i_m} \in \Lambda_o$ with
$\lambda_i = \lambda_{i_1}\lambda_{i_2} \dots \lambda_{i_m}$.

II. Given two d.s.g.l. Λ and M a correspondence θ between the elements of two subsets $\Lambda' \subset \Lambda$ and $M' \subset M$ will be said a translation if it satisfies :

II.1. The correspondence is one to one where ever it is defined.

II.2. If $\lambda_i, \lambda_j \in \Lambda', \theta\lambda_i = \mu_i, \ \theta\lambda_j = \mu_j,$ then $\lambda_i\lambda_j \in \Lambda'$ and $\theta\lambda_i\lambda_j = \mu_i\mu_j$

II.3. The translation will be said :

Total from Λ to M, if $\Lambda' = \Lambda$.
Subtotal from Λ to M, if for all $\lambda_i \in \Lambda$ there is at least a $\lambda_j \in \Lambda$ such as $\lambda_i\lambda_j \in \Lambda'$.

III. A neat coding of Λ into M will be a translation total from Λ to M and subtotal from M to Λ.

In algebraic form we could reduce our axiomatic to :

I' : Λ is the free discrete semi group generated by Λ_o

II' : A translation is an isomorphism between the sub semi groups $\Lambda' \subset \Lambda$ and $M' \subset M$

III': A translation is a neat coding if $\Lambda' = \Lambda$ and M' is a subsemigroup of M neat on the right. (Note that "subsemigroup" entails I.1, I.2 and I.3 ; "free" corresponds to unique in I.4.2 , "discrete" to finite at the same place).

2. Practical significance of the axiomatic :

Let us take a simple example in coding :
$\Lambda_o = \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$; $M_o = \{+, -\}$
(M_o is the usual binary alphabet; Λ is the set of all strings of a finite number of the "elementary messages" λ_i $(i = 1, 2, 3, 4)$ and M is built in the same way with the "letters" + and - .

When coding, we want to establish a correspondence between Λ and some subset M' of M satisfying two conditions :

1) to every $\lambda \in \Lambda$ corresponds at least one $\mu \in M$ ("total" character of the coding)

2) to any distinct $\lambda, \lambda' \in \Lambda$ must correspond distinct $\mu, \mu' \in M'$ in order that the deciphering be free from ambiguity.

A priori any one to one correspondance between Λ and a subset M' from M would do - but usually this could imply that we cannot proceed to the sending of the message before we know it in its totality. So a further practical condition - which is not too easy to formulate rigourously - could be :

3) For a reasonably large number of messages λ the coding is such that for any right multiple λ' of λ (i.e. any $\lambda' = \lambda\lambda''$) the signals μ and μ' have a reasonably long common left divisor μ_1, (i.e. are of the form : $\mu = \mu_1\mu_2$ and $\mu' = \mu_1\mu_2'$).

The simplest way of fulfilling these desiderata is to assign to each $\lambda_{i} \in \Lambda_o$ a string of binary letters μ_i (which very conveniently we may too call a word) and for any sequence $\lambda_{i_1}\lambda_{i_2} \dots \lambda_{i_m}$ to send the corresponding sequence: $\mu_{i_1}\mu_{i_2} \dots \mu_{i_m}$.

For example, with the correspondance : $\mathcal{C}_1$:
$\lambda_1 \to + = \mu_1$; $\lambda_2 \to ++- = \mu_2$; $\lambda_3 \to -+- = \mu_3$;
$\lambda_4 \to -- = \mu_4$
we would have :
$\lambda_2\lambda_3\lambda_1\lambda_1\lambda_2 \to ++--+-++++-$.

It is not obvious however how the set M_o' of the words μ_i has to be selected so that decoding be free from ambiguity :

At my knowledge, the question has been raised first and practically solved by Sardinas and Patterson in a pioneering paper(1).

With the help of semi group concepts we may however obtain a deeper insight into their whole procedure which was purely logical :

We are looking for a total translation from Λ to M and it is quite axiomatic that the decoding is unambiguous if and only if the sub semi group M' generated by M_o' is isomorphoric to the free semigroup Λ - or - for short - that M' is a free subsemigroup of M.

Algebraic consequences of this simple remark are to be found in appendix 1.

Now would come a fourth requirement : (admissibility)

4) The length of the words μ_i must be as small as possible in respect of some a priori probability distribution on Λ.

As a matter of fact (4) will be met incidentally, so to say, in wiew of another condition we put in definition III :

That the translation from $\mathcal{M}$ back to Λ be <u>sub total</u> :

What this means exactly is that any sequence μ of binary digit be a left divisor of at least one message $\mu' \epsilon M'$ which can be completely and exactly retranslated into Λ .

This condition together with the possibility of one-to-one deciphering implies automatically that the code be <u>unitary</u> (as defined below)(see appendix 0), and <u>admissible</u> in that sense that it meets the optimality requirement (4) in respect of at least one a priori probability distribution of the words. (*)

3. <u>Discussion of the decoding methods : scansion</u>

This being settled we have to look more closely at the decoding.

For avoiding repetition let us observe that Λ does not play any role by itself since the $\lambda_i \in \Lambda_0$ are in a one-to-one correspondance with the words $\mu_i \in M'_0$. So we may perfectly well dispense from mentioning it altogether.

But in order to stress when a given string μ of binary symbols is really a set made up of a sequence of words and not any odd sequence of + and - we shall say that μ is a <u>complete message</u> (for instance : " $++--+-$ " $= \mu_2\mu_3$ is a <u>complete message</u>, but " $++--$ " <u>is not</u>) and indicate it by enclosing it into two / signs, which shall denote too, end and beginning of the words.

Let us try to decode the following complete message in code $\mathcal{C}_1$:

$/++--+--+--+/$

The only way open is trial and error : the first + may be:

- either μ_1 itself
- either the first letter from $\mu_2 = /++-/$

so that we have the choice between :

$|+|+--+--+--+|$ and $|++-/-+--+--+/$

In the first case no further doubt comes in and we are lead to :

$|+|+|--|+|--|+|--|+| = \mu_1\mu_1\mu_4\mu_1\mu_4\mu_1\mu_4\mu_1$

In the second we obtain :

$/++-/-+-/-+-/-+ \ = \ \mu_2\mu_3\mu_3 -+$

Since here $-+$ is left at loose end (strictly speaking) the first translation was the good one, being known that the transmission is over. Observe that if, on the contrary, the signal was the same as before except for an added terminal - digit, the conclusion would be exactly opposite :

$|++-/-+-/-+-/-+-/$

is the only fitting "<u>scansion</u>" as we could say by borrowing from prosody this term for its classical flavour.

So the inverse translation from $\mathcal{M}$ back to Λ does not look like satisfying very reasonably the above condition 3.

An obvious remedy to it would be to limit still more the set $\mathcal{M}'_0$. B. Mandelbrot, who has first discussed these problems has distinguished several possibilities :

1) <u>Uniform codes</u> : in which every word has the same <u>length</u> (i.e. number of letters), this criterium <u>giving</u> a direct scansion (examples: all the noise reducing codes introduced so far except for a proposal of "sequential coding" by Peter Elias(4) and some examples by Lemnael(5).)
2) More generally : what we shall call :

   <u>Unitary codes</u> : i.e. codes in which no word is a left divisor of another word (examples : Fano's, Huffman's, Shannon's codes)
3) <u>Natural codes</u> : (introduced by B. Mandelbrot) in which a special letter points out the end of the word (example : most of the spoken or written languages).

Further, Mandelbrot has shown that any unitary code is, at least asymptotically, as good from the point of view of economy of length as any other one. It could seem futile then to care for more extensive classes were we not prompted by other circumstances - and especially by the threat of a noise.

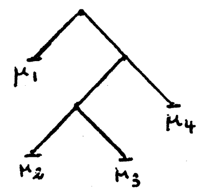
4. <u>Noise absorption and eryodism</u>.

Consider indeed the following code : $\mathcal{C}_2$

$\mu_1 = +;\ \mu_2 = -++;\ \mu_3 = -+-;\ \mu_4 = --$

(which is, parenthetically, just the previous one with the time arrow inverted)

It is <u>unitary</u> all right so that we may represent it by a "tree" in the familiar fashion :



(south west arrow : +

south east arrow : −)

The "<u>neat</u>" condition (<u>subtotality</u> of the translation from $\mathcal{M}$ back to Λ) is reflecting itself in the fact that any branch of the tree

---

(*) If M is the free semi group of all phonemic sequences in English and $\mathcal{M}'$ the sub set of all "<u>semantically correct sentences</u>", $\mathcal{M}'$ is <u>neat</u> in $\mathcal{M}$.

(For instance :

" /pri wat law cut chur coco feet .."(obtained from King Lear, Act III, scene I, with Tippet's help) is fitted into a complete message in $\mathcal{M}'$ by adding : "... and this, Gentlemen, was, may-be, my best example of a semantically void utterance/ ")

ends with a word (for example the code $\mu_1 = +$; $\mu_2 = -++$; $\mu_3 = --$ would not be neat since no word nor sequence of words may begin with /- + - ...).

Suppose that we have to decode the sequence :

/- + - - + - + - - + - -

we obtain directly :

/- + -|- + -|+|- -|+|- -|

and we could have written it down extemporaneously without waiting for the end of the transmission.

But if the first digit had been blurred by noise, this straight forward attitude could not be kept : indeed we decipher the uncertain message

/?+ - - + - + - - + - - ....

either as :

|+|+|- -|+|- + -|- + -|- ....

either as above :

|- + -|- + -|+|- -|+|- -| ....

and as long as the message is going on we have no evidence for deciding between this two interpretations. Things nonetheless are not so bad as they look at first glance :

Suppose that the next letters which appear be + + - + - + - ....

so that up to this time the two alternative versions are :

|+|+|- -|+|- + -|- + -|- + +//- + -|+|- ...
|- + -|- + -|+|- -|+|- -|+|+//- + -|+|- ....

By the seemingly fortuitous fact that in both case the end of a word falls exactly as the same spot (marked // above), the two translations coincidate from this point on and since one of them must be right so is the end of the deciphering - assuming of course that no new error of transmission takes place.

Practically, if such a fact was frequent enough, this would mean that for very low levels of noise, considerable parts of the "meaning" could be preserved. We shall see that this ergodic property (i.e. this relative independence for long sequences of the scansion of the end from that of the beginning) is the rule rather than the exception.

More specifically, for neat codes whose words have all a bounded lenght an apart from three exceptional families there is at least one finite sequence of words - say μ_∞ such that whatever be the initial sequence a, $a\mu_\infty/$ is a complete message. This implies that, when decoding, any blurring or error in a is "absorbed" by μ_∞ and that from the end of μ_∞ on, the scansion starts all right afresh.

Now if the words are given randomly and independently with fixed probabilities, it is clear that the probability for a given sequence not to contain μ_∞ tends with its length exponentially to zero so that any initial error is most likely to have only limited effects.

5. Syntactic equivalence and the fondamental semi groups.

Suppose we be given in code $\mathcal{C}_i$ the following fragment μ from a message :

μ : ... + - - + - - + - + - - ...

By trial and error we see that only three scansions can possibly be fitted to it :

1) ...|+|- -|+|- -|+|- + -|- ...

2) ... +|- -|+|- -|+|- + -|- ...

3) ... + -|- + -|- + -|+|- -| ...

In the same manner the fragment

μ' : ... + - + - + - ...

would give alternatives :

1) ...|+|- + -|+|- ...

2) ... +|- + -|+|- ...

3) ... + -|+|- + -|...

Disregarding the "meaning" of μ and μ' (i.e. their eventual decoding into the Λ language) we may observe that "functionally", so to say, μ and μ' are quite similar :

If the complete message is $/\mu_1 \mu \mu_2/$, the only possibilities are for each of the three scansions :

1) μ_1 is a complete message and μ_2 starts with -/.. or + +/.. or + -/.. (so as to make use of the /- .. left at the end of μ).

2) μ_1 is ending by ../- + (so as to use .. +/) and μ_2 starts as above.

3) μ_1 ends with ../- (for the sake of... + -/) and μ_2 is a complete message.

Easy check shows that the same applies exactly to μ' and we shall say that μ and μ' are syntactically equivalent (*) ($\mu \equiv \mu'$). Actually both are equivalent to an even simpler fragment :

μ'' = ... + - ...

since this last one admits the same scansions :

1) ../+/- .. ; 2) ... +/- .... ; 3) ... + -/...

(*) It is interesting to observe that syntactic equivalence has a direct application to normal linguistics:

If $\mathcal{M}'$ is the set of all sentences grammatically correct :

$\mu_1 \equiv \mu_2$ (approximatively !) if and only if μ_1 and μ_2 pertain to the same grammatical category (for instance in English : both "adjectives" or both "verbs at the third person of the present" etc.)

Now the key point is that for any four finite fragments, μ_1 , μ_2 , μ_3 and μ_4

$\mu_1 \equiv \mu_2$ and $\mu_3 \equiv \mu_4$ implies $\mu_1\mu_3 \equiv \mu_2\mu_4$.

The syntactic equivalence is thus fully compatible with the semi group structure of M and if we consider classes for $\equiv$ (i.e. the subsets of elements from M which are syntactically equivalent between themselves), these classes make a new semi group $\bar{M}$ which is an homomorphic image of M .

$\bar{M} \supset \bar{M}_0$, the fondamental semi group of the coding (f.s.g.) is most usually finite and is easily represented by matrices, but before we explain how, we need still a new concept : that of prefix :

Consider again two fragments μ and μ' but assume, now, that both are beginning at a / mark:

Even if μ and μ' are not syntactically equivalent, it could happen that under this supplementary restriction any further fragment which completes μ into a full message would do the same to μ' :

One could say that " μ and μ' as beginning of messages are syntactically equivalent on the right" (in symbols : $\mu \sim \mu'$)

For example :

$\mu = /--- \ldots$ and $\mu' /+- \cdots$ are not in the relation $\equiv$ (since $/-+\mu'' = /-+-/--/\ldots$ is a complete message although $/-+\mu' = /-++/-\ldots$ is not complete), but $\mu \sim \mu'$ all the same for $\mu\mu''$ is a complete message if and only if

$\mu'' = -/\cdots$ or $++/\cdots$ or $+-/\cdots$

just as well as for μ' .

We call prefixes the classes π_i of fragments for this new relation $\sim$ . .

For the code C_2 , there are three prefixes :

$\pi_1 \ni /\!\!\not{}/$ (words and words only are bringing a $\mu \in \pi_1$ into a complete message.

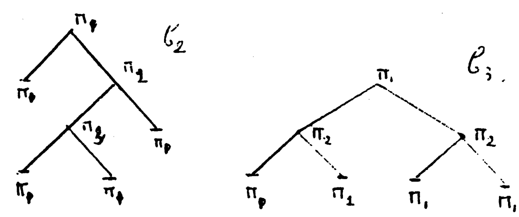
π_1 contains all the words and its existence is typical of unitary coding).

$\pi_2 \ni /-\cdots$ (the corresponding right divisors are $-/\ldots$, $++/\ldots$ and $+-/\cdots$)

$\pi_3 \ni /-+\cdots$ (the corresponding right divisor are $+/\ldots$ or $-/\ldots$).

Now if $\mu_1 \sim \mu_2$, one proves that $\mu_1\mu_3 \sim \mu_2\mu_3$, too, whatever be μ_3

With unitary codes prefixes correspond to nodes of the tree in a one to many fashion : Two nodes being in relation $\sim$ ("pertain to the same prefix") if the subtrees below them are identical. Such things does not occur in our C_2 code (see below), but are quite typical of uniform codes.



In the code C_3 of length 2 ($\mu_1 = ++$; $\mu_2 = +-$; $\mu_3 = -+$; $\mu_4 = --$) there is only two prefixes : one, π_1 , corresponding to complete messages - i.e. to sequences with an even number of letters - and another one, π_2 , corresponding to odd length sequences.

6. Matrix representation of the fundamental semi group.

If we have started reading just at the beginning of the transmission, we may consider at any time t the prefix $\pi(t)$ to which pertain the initial fragment till the t-th letter as a "state" which changes at any new letter received.

For instance - apart from any meaning again - the sequence $/+--+--\ldots$ corresponds to the following sequence of prefixes :

$\pi_1\,\pi_1\,\pi_2\,\pi_1\,\pi_2\,\pi_3\,\pi_1\,\pi_1\,\pi_2 \;\cdots$

It is easy to visualise "+" and "-" respectively as the transition matrices :

$$\begin{array}{c|ccc|} & \pi_1 & \pi_2 & \pi_3 \\ \hline \pi_1 & 1 & 0 & 0 \\ \pi_2 & 0 & 0 & 1 \\ \pi_3 & 1 & 0 & 0 \\ \hline \end{array} \qquad \begin{array}{c|ccc|} & \pi_1 & \pi_2 & \pi_3 \\ \hline \pi_1 & 0 & 1 & 0 \\ \pi_2 & 1 & 0 & 0 \\ \pi_3 & 1 & 0 & 0 \\ \hline \end{array}$$

(+) (-)

(+ lets π_1 invariant since it is a word. It sends π_2 into π_3 and makes a word from π_3 etc..)

These matrices correspond in a one to one fashion to the elements of the fondament semi group, for instance:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

(+ +) (- +)

(with the usual line by column multiplication) is the matrice given below. What are the matrices corresponding to complete messages ? In the general case they are the matrices of the subsemigroups $\overline{M}_1'$ image of M' by the syntactic homomorphism .

But if the code is unitary, $\overline{M}'$ is characterised very nicely, since $\mu \in M'$ implies that μ sends π_1 into itself : $\overline{M}'$ is just the set of the matrices of with 1 in the top left corner.

Further, noise absorption - or ergodic - properties reflect themselves quite directly on this matrix representation.

Suppose that the correct message $/\mu\ldots$ and the perturbated message $/\mu'\ldots$ fall back both at this very time on a common scansion mark $/$. If the prefix corresponding to μ was π and that corresponding to μ' was π' , this would mean that the next signals sends both π and π' into $\cdot\pi_1$.

On the matrices this is expressed by the fact that in column π_1 there is two 1's : one in the line π and another one in line π' . In particular μ_∞ is a matrix with 1 everywhere in column π_1 .

But this in turn is linked closely with the fact that $\overline{M}$ is a <u>semi</u> group and not a group (whose matrices should all have a single 1 by column).

Consider as a counter example the uniform code with four words :

Its f.s.g. is just the cyclic group of order two, made up of the two elements :

$$\begin{array}{c|cc|} & \pi_1 & \pi_2 \\ \pi_1 & 0 & 1 \\ \pi_2 & 1 & 0 \end{array} \quad \begin{array}{c|cc|} & \pi_1 & \pi_2 \\ \pi_1 & 1 & 0 \\ \pi_2 & 0 & 1 \end{array}$$

(+ or - or any <u>odd</u> length sequence)

($\emptyset$ or ++ or +- or ... etc. or any <u>even</u> length sequence).

No real absorption takes place for indeed if we had missed the first letter of the transmission and started wrongly scanding from the second letter, the error will obvious go on as long as does the message.

As a matter of fact uniform codes are the only neat codes with a bounded length for words whose f.s.g. is a group. They are the first exceptional non ergodic family.

7. <u>Super coding</u>.

We have given a very general definition of "translation" which suggests the possibility of more complex processes involving not only <u>two</u> but several languages. In the general case, things are a bit confused and we shall restrict ourself to <u>Unitary Neat</u> Coding from K into Λ and from Λ into M .

Suppose for instance that we have the following set up :

K is a d.s.g.l. with words K_i $(1 \leq i \leq 7)$

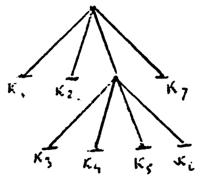
Λ is a d.s.g.l. with words λ_i $(1 \leq i \leq 4)$

M is our familiar binary d.s.g.l.

Each word of Λ is coded in M as in example 2:
$\lambda_1 \to +$; $\lambda_2 \to -++$; $\lambda_3 = -+-$; $\lambda_4 = ---$.

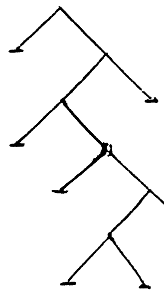
Each word of K is coded by the following sequences $\lambda^{(j)}$ of Λ (for clarity we use upper and lower indices) : $K_1 \to \lambda^1 = \lambda_1$; $K_2 \to \lambda^2 = \lambda_2$; $K_3 \to \lambda^3 = \lambda_3\lambda_1$
$K_4 \to \lambda^4 = \lambda_3\lambda_2$; $K_5 \to \lambda^5 = \lambda_3\lambda_3$; $K_6 \to \lambda^6 = \lambda_3\lambda_4$; $K_7 \to \lambda^7 = \lambda_4$.

This coding is unitary and neat all right and corresponds to the tree :



Now there is again a coding of K into M when every λ^j is written in binary alphabet :
$K_1 \to +$; $K_2 \to -++$. $K_3 \to -+-+$; $K_4 \to -+--++$;
$K_5 \to -+--+-$; $K_6 \to -+----$, $K_7 \to -+----$.

It is not difficult to see that this $K \to M$ coding is unitary and neat. Its tree is given below.



Since we know the importance of fundamental semi groups we would be interested to get at once that ($\overline{\overline{M}}$) of the $K \to M$ process from the other two ($\overline{\Lambda}$ for $K \to \Lambda$ and $\overline{M}$ for $\Lambda \to M$) or, alternatively, to know the relation between the syntactic equivalences on the bottom structure M . $\equiv (\Lambda$ in respect of $\Lambda \to M$, without K appearing in the picture and $\equiv (K)$ in respect of $K \to M$ with Λ put off from the circuit.

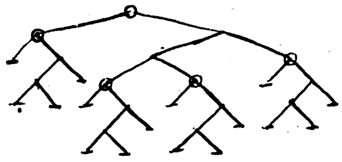
The main result is that :

$\mu_1 \not\equiv \mu_2\ (\Lambda)$ entails $\mu_1 \not\equiv \mu_2\ (M)$
or, if one prefers, <u>that $\overline{M}$ is a homomorphic image of $\overline{\overline{M}}$</u> .

This is rather convenient from a technical point of view for it allows what is called a filtering. If starting from the assumption that the λ_i are provided independently with fixed probabilities by the source, we discover later on that, actually, they were just building blocks in some higher degree semantic units (sent again independently of each other as a second approximation) we can preserve at least some of the features of our initial approximation.

But the main point for us here lies in another aspect.

Suppose that the $K \to \Lambda$ coding be uniform. in general the $K \to M$ one will not be so, but it will fail to be ergodic just the same, giving us the second of the three exceptional families mentionned above. We shall call such codes "uniformily composed codes". An example is given below :



$\mathcal{C}_2 \otimes \mathcal{C}_2$

$K \to M$

$K \to \Lambda$: uniform of length two

$\Lambda \to M$: our usual $\mathcal{C}_2$

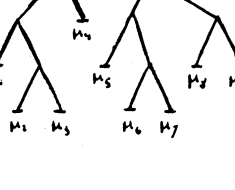
The nodes indicated with a o are the ones corresponding to nodes in the $K \to \Lambda$ coding.

8. Anagrammatic codes.

Let us come now into the last family. For this we produce the following horrible example : $\mathcal{C}_4$

$\mu_1 = +++;\ \mu_2 = ++-+;\ \mu_3 = ++--;\ \mu_4 = +-.$

$\mu_5 = -++,\ \mu_6 = -+-+;\ \mu_7 = -+--,\ \mu_8 = --+,\ \mu_9 = ---.$



$\mathcal{C}_4$ is not uniform - nor composed uniformily of a smaller code. But it has the property that by inverting its words we found again a unitary code and, indeed, its symmetric image (symmetric in respect of the N.S. line !)

Since ergodicity is somewhat synonimous of irreversibility of time, we are put on the alert by this oddity.

Indeed, absorption is linked very closely with the problem of reading "backward" messages with an inverted code, but, without entering this amusing theory, we can see at once that $\mathcal{C}_4$ and all its family are not ergodic.

If a code is unitary the only sequence which let π_1 invariant are the complere messages, whose set is M'. In symbols, this means :

$$\mu_1\mu_2 \in M' \quad \text{and} \quad \mu_1 \in M' \quad \text{implies} \quad \mu_2 \in M'$$

Suppose now that the same property be true on the other direction, i.e. that we had :

$$\mu_1\mu_2 \in M' \quad \text{and} \quad \mu_2 \in M' \quad \text{implies} \quad \mu_1 \in M'$$

Let μ_1 be a complete message which is the unperturbated beginning of the transmission; μ'_1, its noise corrupted form and μ_2 any other complete message. By the above condition $\mu'_1\mu_2$ may have a final scansion like that of $\mu_1\mu_2$ if and only if μ'_1 is a complete message, too.

As this is usually not the case the error will go on till the end.

Codes which are unitary for both directions of time (anagrammatic codes) are not yet fully explored but a construction for various infinite families of them is known. With binary alphabet, there is just the one given above and its symmetric for less than 16 words. It is conjectured that there is still no more than 38 other one below 32 words (on about 10^{16} distinct usual unitary neat codes of this size or less!).

So the family is really exceptionally interesting and deserves further studies since with the uniform and the uniformily composed codes, anagrammatic codes are the only length-bounded codes escaping ergodicity.

References.

1. A.Sardinas and Patterson (1953) Convention records of the I.R.E.
2. B.Mandelbrot .
   1953. Proc.Symp.Comm.Theory. London.
   1954. Proc. Symp. Inf. Network.
   1955. Proc. Symp. Comm. Theory . London.
3. P. Dubreil.
   1941. Mem. Acad. Sci. p. 1-52.
   1951. Rendiconti di Math. 81 p. 289 - 306.
   1953. Bull. Soc. Math. (10) p. 183 - 200.
4. P.Elias.
   1955. Proc. Symp. Comm. Theory. London.
5. A.E.Laemmel.
   1953. ibid.