INFORMATION AND CONTROL 6, 331–336 (1963)

# A Simple Proof of the Decipherability Criterion of Sardinas and Patterson

G. Bandyopadhyay

*Department of Mathematics, Indian Institute of Technology, Kharagpur, India*

A simple alternative proof is given for a necessary and sufficient condition for the decodability of a sequence of codes. The proof involves no application of linear syntactical bases or other sophisticated algebraic criteria.

A necessary and sufficient condition for unique decipherability of coded messages has been obtained by Sardinas and Patterson (1953). The condition is simple, but the proof requires quite sophisticated methods. The formalism necessary for the proof, as described by the authors (*loc. cit.* para. 2, p. 106), needs the knowledge of semi-groups with some special properties (the "linear syntactical basis"). The formalism has been given partly in the reference given above and partly in a short abstract (Patterson, 1951). The complete proof appears in an unpublished research report (Sardinas and Patterson, 1950).

Since the condition is important in Information Theory a simplified proof without invoking any sophisticated formalism may be thought desirable. The desirability is further enhanced due to the fact that "linear syntactical basis" has not been necessary so far in any branch of Information Theory except for the above proof. For completeness the condition is restated and the terminology clarified in the next section before proving it.

## TERMINOLOGY AND THE STATEMENT OF THE THEOREM

We are given a set of *code symbols* (called symbol by Shannon (1948), coding digit by Huffman (1952), letters by Gilbert and Moore (1959), and code letter by Karp (1961)), like dot or dash in morse code. We shall designate them as $D$, $E$, $F$, etc. (the symbols $A$, $C$ being reserved for other purposes). From these symbols any finite *symbol sequence* like *DE*, *D*, *EFD*, etc. can be constructed. Of these sequences some are selected to represent messages. Each of these will be called *codes*, following Shannon (1948) and will be denoted by $C_1$, $C_2$, $\cdots$ ; $C_1'$, $C_2'$, $C_3'$, $\cdots$ , etc. A sequence of codes will be called a *code sequence*. Thus if *D*, *DE*, *ED* can be codes then *DDDDEED* is a code sequence, but *EEEEE* is not a code sequence. A code sequence is a symbol sequence; but a symbol sequence is not always a code sequence. The aggregate of the set of all codes will be called a *code set* (called encoding by Gilbert and Moore (1959) and code by Karp (1961)). Following Huffman (1952) one symbol sequence will be called the *prefix* of another if the former can be obtained by removing some symbols from the end of the latter. The symbol sequence left over in the latter by taking away the prefix will be called, following Sardinas and Patterson (1953), a *member of segment 1* constructed from the code set.

*Segment 1* is complete when all its members have been obtained by considering each member of the code set and examining if it is a prefix of another. If now we take each member of segment 1 and examine if it is a prefix of any code or if any code is its prefix then the symbol sequence obtained by removing the prefix is called a *member of segment 2* and the totality of members constitute *segment 2*. Segments 3, 4, 5, $\cdots$ are constructed from segments 2, 3, 4, $\cdots$ respectively, just in the same way as segment 2 was constructed from segment 1. Evidently a code set is *decipherable* (i.e., any code sequence can be read unambiguously) if no two distinct code sequences have the identical symbol sequence. The theorem of Sardinas and Patterson states: *A Code Set is decipherable if and only if no segment contains a Code.* We proceed to give a simple proof of the theorem in the next section. We shall have to use the idea of (geometrical) segment of a line in the proof. Unless qualified by the adjective "geometrical" or marked by the endpoints, e.g., *AA*, etc. the word segment will have the meaning as in the definitions above. Additional symbols will be used, which will be explained where they are needed, viz., at the latter parts of the next section.

## PROOF OF THE THEOREM

To prove that the condition is necessary let us suppose that two code sequences, $C_1$, $C_2$, $\cdots$ , $C_k$ and $C_1'$, $C_2'$, $\cdots$ , $C_{k'}'$ have identical symbol sequences. Let us further assume, without any loss of generality, that no two codes other than $C_k$, $C_k'$ in the above sequences end simultaneously. For convenience of our argument we shall represent the codes by (geometrical) segments of a straight line in Fig. 1. Let $A$ mark the beginning of $C_1$ and $C_1'$. Let $A_1$, $A_2$ mark the end of $C_1$, $C_2$ and also the beginning of $C_2$, $C_3$ respectively. Thus if $C_1$ stands for $DEFD$ and $C_2$ for $FED$ then, $AA_1$ will mean $DEFD$ and $A_1A_2$ will mean $FED$. Let $A_1'$, $A_2'$, $\cdots$, $A_{k'}'$, have similar meanings in relation to $C_1'$, $C_2'$, $\cdots$ $C_{k'}'$. If $C_1'$ stands for $DE$ then $A_1'A_1$ will mean $FD$. We shall prove that at least one member of one of the segments is a code.

Since $AA_1$ and $AA_1'$ are codes it follows that $A_1'A_1$ is a member of segment 1. (This is provided $A_1'$ lies to the left of $A_1$, which we assume. Had it been otherwise the corresponding argument would run analogously.) The member of segment 2 that can be constructed from $A_1'A_2'$ is either $A_2'A_1$ or $A_1A_2'$, depending on whether $A_2'$ is to the left or right of $A_1$. Construction of segments will thus proceed and the $r$th segment will be $A_1A_r'$ where $A_r'$ is the first of the $A'$'s lying to the right of $A_1$. (It may be noted that where $A_2'$ is to the left of $A_1$, $A_2'A_1$ is obtained from the first segment-member $A_1'A_1$ by taking away from it the code $A_1'A_2'$ as prefix. In obtaining $A_1A_r'$, however, the segment-member $A_{r-1}'A_1$ has been taken away as prefix from the code $A_{r-1}'A_r'$). Continuing in this manner it follows that $A_r'A_l$ is a member of $(r + l - 1)$th segment where $A_l$ is the first of $A$'s lying to the right of $A_r'$. Thus every mark, for some finitely numbered segment, becomes either the beginning or the end of a member of a segment. Further, every segment has its beginning in one of the $A$'s and end in one of the $A'$'s or vice versa. (This statement can be proved by induction, for if $A_r'A_l$ be a segment then the next segment is $A_lA_{r+1}'$ or $A_{r+1}'A_l$ according as $A_r'A_l$ is the prefix of a code or has a code as prefix. Also the member $A_1'A_1$ or $A_1A_1'$ is of the required form.) Now let $A_{k-1}$ be the mark nearest to the common

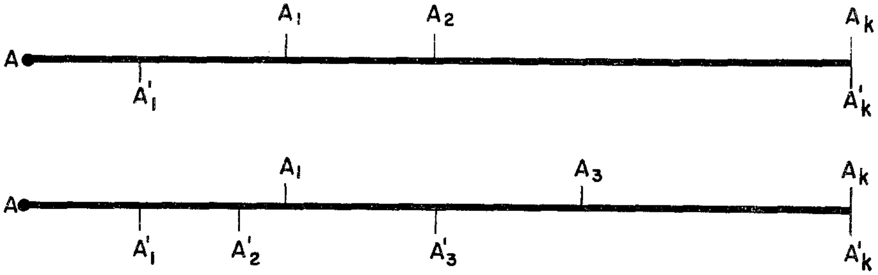

FIG. 1. Representation of codes on a line segment

point $A_k$ , $A'_{k'}$ . If $A_{k-1}$ is the end of a segment, the segment must be $A'_{k'-1}A_{k-1}$ . Hence the next segment, found by taking this away as prefix from the code $A'_{k'-1}A'_{k'}$ is $A_{k-1}A'_{k'}$ , i.e., $A_{k-1}A_k$ which is a code. Similar arguments could be forwarded if $A'_{k'-1}$ would have been nearest to $A_k$ . This completes the proof of necessity.

To prove the sufficiency we assume for definiteness that a member of segment 3 is a code. We shall demonstrate how two code sequences can be constructed which have an identical symbol sequence. Figure 2 shows different possibilities (i), (ii), (iii), (iv) that arise when a member of segment 3 is a code. Figure 3 shows two code sequences with identical symbol sequence that can be constructed in cases (i), (ii), (iii), (iv). We describe below the essence of this construction.

To construct two alternative code sequences each step of Fig. 2 is written down as purely algebraic equations as shown below. The codes are then transposed so that no negative sign remains on either side. The set of codes on each side, taken in order, in which they occur in Fig. 2, represents the two alternative code sequences. To illustrate this, consider case (iii). The equations are:

$$S_p^3 = C_p = C_q - S_q^2 \quad \text{or} \quad S_q^2 + C_p = C_q \qquad (1)$$

$$S_q^2 = S_r^1 - C_r \quad \text{or} \quad C_r + S_q^2 = S_r^1 \qquad (2)$$

$$S_r^1 = C_s - C_t \quad \text{or} \quad C_t + S_r^1 = C_s$$

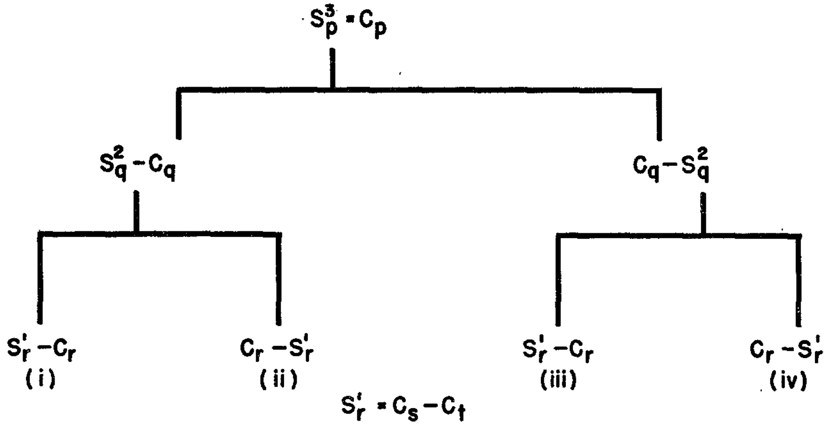

FIG. 2. $S_p^k$ stands for the $p$th member of segment 1. $S_r^n - C_r = S_q^{n+1}$ stands for the fact that $S_q^{n+1}$ has been formed by removing prefix $C_r$ from $S_r^n$. $C_r - S_r^n = S_q^{n+1}$ stands for the fact that $S_q^{n+1}$ has been formed by removing prefix $S_r^n$ for $C_r$ .

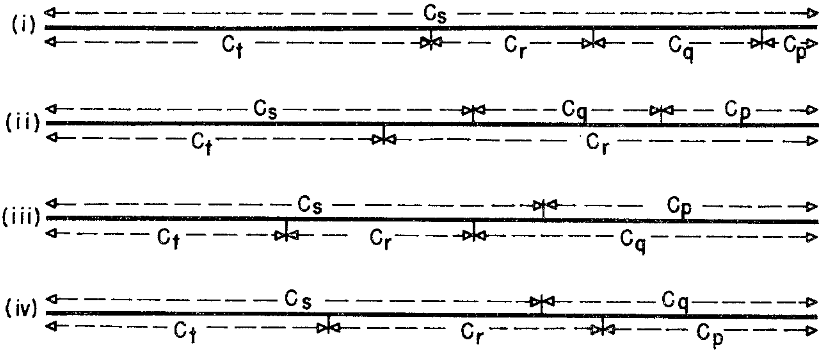

FIG. 3. Code sequences written above the line give one decomposition and the ones written below give another decomposition.

Eliminating $S$'s we have

$$C_p + C_s = C_t + C_r + C_q \tag{3}$$

We find the correspondence of this equation in Fig. 3(iii) where $C_s$ followed by $C_p$, written above the line, has the same code sequence as $C_t$ followed by $C_r$ followed by $C_q$, written below the line.

That the above process is justified for Eq. (1) can be taken to mean that $S_q^2$ followed by $C_p$ has the same symbol sequence as $C_q$. Placing now $C_r$ to the left of each and using (2) one gets

$$S_r^1 + C_p = C_r + C_q$$

meaning that $S_r^1$ followed by $C_p$ has the same symbol sequence as $C_r$ followed by $C_q$. Proceeding in this fashion Eq. (3), with analogous meaning, follows.

All this means that we can handle (1), (2), etc. as algebraic equations provided proper order is maintained. The method can similarly be extended when a member of any other segment say 4, 5, 6, $\cdots$ is a code. This completes the proof.

The literature on the subject (see references) contains some examples and also clarifies the interrelation between this condition, the prefix condition, and the usual condition of having a fixed terminating symbol. We do not enter into the discussion of these in this paper.

## Acknowledgment

The author wishes to warmly thank Dr. R. B. Banerji for taking considerable care in going through the composition of the paper.

Received: November 14, 1962

## References

Gilbert, E. N. and Moore, E. F. (1959), Variable length binary encoding. *Bell System Tech. J.* **38,** 933–968.

Huffman, D. A. (1952), A method for construction of minimum redundancy codes. *Proc. I.R.E.* **40,** 1098–1101.

Karp, R. M. (1961), Minimum redundancy coding for discrete noiseless channels. *Trans. I.R.E.* **IT-7,** 27–38.

Patterson, G. W. (1951), Linear syntactical bases. *J. Symbolic Logic* **16,** 236.

Sardinas, A. A. and Patterson, G. W. (1953), A necessary and sufficient condition for the unique decomposition of code messages. *IRE Intern. Conv. Record* **8,** 104–108.

Sardinas, A. A. and Patterson, G. W. (1950), A necessary and sufficient condition for the unique decomposition of codes messages, Research Division Report, 50-27, Moore School of Electrical Engineering, University of Pennsylvania.

Shannon, C. E. (1948), A mathematical theory of communication. *Bell System Tech. J.* **27,** 379–423, 623–656.