

**SYSTEM FOR CONVERTING MEDICAL
INFORMATION INTO REPRESENTATIVE
ABBREVIATED CODES WITH
CORRECTION CAPABILITY**

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to coding systems and more particularly but not solely to coding algorithms which express ideas or information as codes for later use.

2. Description of the Prior Art

Most existing systems which express ideas or information as codes deal with limited information, for example, a name or a stock item. The connection between the code and the information is rigid and the method of assessing the data is uniform. In many industries and professions, however, this degree of uniformity is not present. In medicine, for example, there are tens of thousands of words which legitimately expressly subtly different concepts and numerous similes and similar sounding words with quite different meanings, while totally dissimilar phrases may express very similar information.

Existing systems which translate medical information (including diagnoses and operations) into computer codes include the World Health Organization's ICD9 system and the British READ system. ICD9 is a numeric coding system where diagnoses and/or operations are expressed as a floating point number. For example, "812.3" in the ICD9 system is a "Fracture of humerus" (upper arm bone). Systems which rely on the ICD9 coding system typically employ clerical coders to examine the clinical notes and then find the appropriate code which matches the information in the clinical notes.

The READ system uses an alphanumeric code, with combinations of letters and digits used to express the information. Unlike the ICD9 system, the READ system is hierarchical in that diagnoses and operations with increasing degrees of complexity retain the same stem part of the code with additional information added to the end. Clerical coders are commonly employed to manually translate surgeons' notes into the READ codes. The clerical coders achieve this by selecting words and similes from tables of commonly used terms which are also referenced to the appropriate codes. In both the ICD9 and READ systems there are many opportunities for errors to make their way into the coding system. For example, the coders are usually not highly trained in the field which they are coding and are, therefore, not uniformly reliable to the degree of accuracy required if medical decisions are to be made from the coded information.

Extraction or interrogation of information using existing coding systems also requires a sound knowledge of the subject. In existing medical coding systems, for example, very similar diagnoses may have totally dissimilar codes. An example is Rheumatoid arthritis which has ICD9 code 714 while Charcot's Arthritis has ICD9 code 250.6. A search through coded information for "arthritis" will, therefore, need to include all such widely varying codes. Similarly, "Monteggia's Fracture" is actually a combination of a fracture and a dislocation or subluxation. Therefore, any searches for "injured joints" will need to include Monteggia's fracture, Galeazzi fracture and a host of others whose names and codes give the reader no clue to their nature. This non-uniformity of codes requires a considerable degree of sophistication and vigilance by researchers.

SUMMARY OF THE INVENTION

It is, therefore, an object of the present invention to provide a coding system which goes some way towards

overcoming the above disadvantages or which at least provides workers in the field with a useful choice.

Accordingly in one aspect the invention consists in a method of coding data in which codes are assigned to the data. The method comprises the steps of:

- i) Dividing the data into sub-parts, each said sub-part being of a particular sub-part type.
- ii) Preliminarily grouping the sub-parts into sub-groups by introducing divisions in the data. Each sub-part in a particular sub-group sharing an association with all other sub-parts in that particular sub-group.
- iii) Examining each of the sub-group for completeness in regard to each sub-group containing a predetermined distribution of sub-part types.
- iv) Completing the sub-groups by amending the constituent sub-parts so that each sub-group contains the predetermined distribution of sub-part types.

In a second aspect the invention consists in a program, for execution by a computer, which analyses input data describing an event, item or operation and produces coded output to represent the input data and also attempts to correct deficiencies in the data. The program comprises the steps of:

- i) Dividing said input data into sub-parts, each sub-part being of a particular sub-part type.
- ii) Preliminarily grouping the sub-parts into sub-groups by introducing divisions in the data. Each sub-part in a particular sub-group sharing an association with all other sub-parts in the particular sub-group.
- iii) Checking each said sub-group for completeness in regard to each sub-group containing a predetermined distribution of sub-part types.
- iv) Attempting to complete the sub-groups so that each sub-group contains the predetermined distribution of sub-part types by analysing the sub-groups in relation to each other and carrying out completion operations on the sub-parts and sub-groups.

The invention consists in the foregoing and also envisages constructions of which the following gives examples.

BRIEF DESCRIPTION OF THE DRAWINGS

One preferred form of the present invention will now be described with reference to the accompanying drawings in which;

FIG. 1 is a flow diagram showing a broad overview of the coding algorithm of the present invention;

FIG. 2 and FIG. 3 are consecutive flow diagrams, detailing the coding system of FIG. 1;

FIG. 4 is a flow diagram expanding on the Lexicon application block of the flow diagram of FIG. 3; and

FIG. 5 is a highly stylized diagram of a human with areas of the body represented by example code snippets used in the flow diagram of FIG. 1.

**DETAILED DESCRIPTION OF THE
PREFERRED EMBODIMENT**

Broad Overview

In the figures each rectangular block represents an algorithm. With reference to FIG. 1, a flow diagram broadly illustrating the method of coding of the present invention is shown. The flow diagram of FIG. 1 may be, for example, carried out in the execution of a computer program. An example of a high level computer language which may be used to write a program to facilitate execution of the coding