

An encoding space designed for application in encoding localizable sentences for communication through the language barrier

William J. G. Overington

Thursday 10 August 2017

This space, which I choose to name star space, is entered by a character sequence consisting of an integral sign character followed by three circled digit characters. The space is thereafter exited in certain circumstances.

1. If the first and third circled digit characters are the same yet not a circled zero, then exit is after as many additional circled digit characters as indicated by the third circled digit character.
2. If the third circled digit character is a circled zero, then, regardless of what is the first circled digit character, the next three circled digit characters after the third circled digit character taken together determine how many more circled digit characters there are after those next three circled digit characters.
3. Otherwise, exit is after the third circled digit.
4. In expressing star space within an information technology application, if any character other than a circled digit, a space, a comma, a carriage return, a line feed is encountered before when exit would be expected then exit takes place as an error exception in the expressing of the coding in the particular situation.

Thus many sequences in star space consist of an integral sign and three and only three circled digit characters. Certainly, the way that star space is designed could result in some valid sequences of over one thousand circled digit characters, for example, if a sequence were to start with, for example, ∫②⑤①②③④ as nine hundred and ninety-eight more circled digit characters would then be needed to complete the sequence. There is no present intention to use such long sequences and at present there is no known application where that would be desirable. The format is simply chosen so as to minimize the possibility of future developments running out of encoding space.

* * *

The star space is designed for application in encoding localizable sentences for communication through the language barrier.

In application, not all of the sequences available within the space are used.

In particular, any star space sequence that has two consecutive characters the same as each other is not used for encoding a localizable sentence. This is based upon the experience of the present author who at one time had a transaction card with a repeated digit in the number and found that there was often a problem in communicating the

number verbally over a telephone link due to there being confusion over whether one digit was being repeated or whether there were two digits the same one after the other.

Level 1.

In level 1 application, a circled zero is not used. This is because in some circumstances the sequence may need to be spoken using just digits, for example over a telephone link. The words for digits are often taught when a foreign language is taught, yet often zero is not taught. So a person knowing only nine digits in a language could potentially use this system over a telephone link so as to communicate through the language barrier to some extent. Also communication could be by using plastic numbers on a desk, and sets of plastic numbers do not always include a zero.

Also, in level 1, no digit is used twice in any one sequence. Thus there are five hundred and four (equals nine times eight times seven) sequences used in level 1. These are intended mostly for application in communication within medical situations, including everyday communication between a patient and a member of staff in a hospital or a care home. For example, "Would you like a drink of water?" and "Would you like to go to the day room?". For specific meetings a human translator is often arranged: this localizable sentence system is intended for everyday communication and for communication in emergencies. However, some of these five hundred and four sequences are for more general communication, such as "Good day." and "Best regards,".

The encoding does include some extremely long sequences. There is no present intention to use very long sequences, yet these are designed into the encoding space so that if people write software to detect star space sequences in text streams in information technology applications, then if some future development needs a longer sequence then the software will accept that sequence as a valid sequence for analysis.

Please consider the following example of a star space sequence together with its meaning localized into English.

∫⑦⑧① Would you like a drink of water?

As GS1-128 bar code technology is being introduced into National Health Service hospitals in the United Kingdom, a specific format is included in this present document for expressing a star space sequence within a GS1-128 barcode.

GS1-128 barcodes can contain characters, though only from a limited set of characters.

For this research, and maybe for long term practical application, the following format is designated for use of star space sequences within GS1-128 barcodes.

Instead of the integral sign use a sequence of two colon characters.

Instead of circled digits use ordinary digits.

At the end add a colon and a semicolon.

Use GS1-128 Application Identifier 97. That is effectively a Private Use code for use within an organization.

Thus ∫⑦⑧④ in star space is, expressed in the way that GS1-128 encoding is expressed in printed text that accompanies the barcode, suggested to be as follows.

(97)::781.;

The internal encoding method of the barcode is not quite like that, but that is the format of the text that accompanies the barcode.

* * *

Beyond Level 1.

An example of the application of sequences that use more than three circled digit characters is to use sequences that start with ∫③①③ so as to encode a collection of localizable sentences relating to enquiring through the language barrier about relatives and friends after a disaster using email, possibly, and desirably, with automated localization in the computer of the recipient of the message. Quite detailed sentences could be used: for example, "The enquirer is the brother of the first person that was named."

Also, longer sequences could be used, specified using a sequence within a very sparsely encoded part of star space, so as to ensure that some messages can only be conveyed by a very explicit encoding and could not be received in error by just a one bit error, or a several bit error, during the transmission of the star space sequence for a localizable sentence.

* * *

Here are sequences and meanings localized into English of the sequences thus far assigned in star space in the research project. These sentences are intended as examples so as to provide a catalyst for thought of what could possibly be achieved with a grater number of encoded localizable sentences.

∫①②③ Good day.

∫②⑤① Have you vomited?

∫②⑤③ I have vomited.

∫②⑤④ I have not vomited.

∫381 Do you have pain?

∫382 I have pain.

∫384 I do not have pain.

∫385 I have pain always.

∫386 I have pain sometimes.

∫387 I have pain, but it is not very much.

∫389

∫391 Where do you have pain?

∫573 Shall we use system 'five seven three'?

∫781 Would you like a drink of water?

∫782 I would like a drink of water.

∫783 I would not like a drink of water.

∫812 Would you like to go to the day room?

∫813 I would like to go to the day room.

∫814 I would not like to go to the day room.

∫815 Would you like to go to your bed?

∫816 I would like to go to my bed.

∫817 I would not like to go to my bed.

∫987 Best regards,