CodeFacts
AIDC (Automated Identification and Data Collection) Technical & Informational Documents
Written for Everyone

Bar Code Symbologies

A bar code symbology is much like an alphabet is to a language. A given combination of bars and spaces represent a particular character, much the same way that “dot-dot-dot” represents the letter 'S' in Morse Code. Think of bar code symbols as being made of characters from a particularly strange typestyle or font—a *machine-readable* font, as opposed to the typical human-readable.

Some bar code symbologies have two module widths (i.e., bar and space thicknesses) and some have four. As with most bar codes, all can be shrunk and expanded to some limit as long as the relationship between the two or four module widths remains consistent within the symbol.

Code 39 (Code 3 of 9)

Synopsis: Code 39 was formerly the most popular industrial-use code in the world. It is a variable-length code which allows numbers, upper-case alpha characters and seven other punctuation characters – - . $ / + % and space.

Technical Details: There are two bar and space widths in Code 39 (i.e., fat and thin bars). The ratio between the elements can be anywhere from 2:1 to 3:1. The latter is preferred, although this makes the symbol longer in length. There is a special character, designated as an asterisk (*) which is used as both the start and stop character. The reliability of Code 39 is one of the highest, with misreads seldom ever occurring. Rarely, a single Modulus 43 check digit is added to the end of the symbol in order to check its validity.

Advantages: Most any bar code reader will read Code 39. It is a flexible symbology because of its ability to encode virtually any number of alpha as well as numeric characters within a single symbol.

Disadvantages: Because Code 39 uses only two module widths, it is not a very dense code (i.e., it takes many elements to make up a single character, hence a

larger symbol). If space for the symbol is at a premium, a Code 39 symbol may not fit.

Uses and Industries: Many industries (e.g., automotive, health care, US. Dept. of Defence) have adopted Code 39 as a standard.

Full ASCII Code 39 (Extended Code 39)

Synopsis: This code allows all 128 lower-ASCII characters to be encoded in a bar code symbol but is not recommended for general use.

Technical Details: Invented to make up for some of the misgivings of Code 39, this symbology really has little to offer. In order to encode lower-case alpha, punctuation and other ASCII code characters, it uses doubled-up Code 39 bar code characters to create characters not found in the normal Code 39 set. For example, to encode a lower-case 'e', an upper-case 'E' is followed by a '+' character. In essence, in order to extend the character set, you could end up making the symbol much longer and Code 39 is not a very space-efficient bar code to begin with.

Advantages: Has a large allowable character set and can be easily printed with software for printing regular Code 39 characters.

Disadvantages: Takes up piles of space and really offers nothing compared to symbologies such as Code 128. While most bar code readers will read Code 39, many will not properly interpret the Extended Code 39 set.

Uses and Industries: Almost none.

Codabar

Synopsis: Essentially a numeric-only bar code symbology of variable length but is not very space-efficient. Consumes virtually the same amount of space as Code 39.

Technical Details: Like Code 39, Codabar uses two element widths and is variable in length. There are four characters designated as 'start' and 'stop' characters, one of which must appear at each end of the symbol. They are designated by the letters a, b, c and d, although no alpha characters are allowed in Codabar.

Advantages: Fairly reliable. Can be printed using rotating sequential numbering heads on business forms presses.

Disadvantages: Not very space-efficient and limiting with only numeric characters allowable.

Uses and Industries: Libraries, blood banking, courier package tracking.

Code 128

Synopsis: The best and most compact bar code currently available. Allows for great flexibility in encoding all 128 lower-ASCII characters into a variable-length symbol.

Technical Details: Code 128 uses four module widths to represent each character in the symbol. This allows for very efficient use of lineal space. There are three "character sets" in Code 128. This can be confusing but can be explained:

The three character sets in Code 128 as designated as A, B and C. Set A contains numbers, upper-case letters, punctuation and most of the 'non-printing' ASCII control characters. Set B contains both upper- and lower-case letters, numbers and punctuation. Finally, set C basically contains all of the two digit characters from 00 through 99.

A particular unique pattern of bars could represent a character in any one of the sets. For instance, there is a particular bar-space pattern that represents the character 'BELL' (ASCII 7) in Set A, a lower-case 'g' in Set B and the number '71' in Set C. How does the bar code reader know which character(s) of the three that unique pattern of bars represents? There are another set of characters called "shift" characters, which are kind of like the shift key on a keyboard. A 'Shift A' character precedes all characters to be interpreted in that characters set and so on with 'Shift B' and 'Shift C'.

With these shift characters and three different character sets, it gets very hard to predict just how long a bar code will be because of the number of shift characters that may or may not appear in the symbol. Obviously, sticking to one of the three sets for the whole symbol would seem to be the best but because of the added benefits of Set C, this is not always so.

Code 128 Set C is referred to as *double-density numeric* because each bar code character can represent *two* numbers! This is far more efficient than the usual situation of having one character represent one number. For example, if you

wanted to incorporate the characters ‘ABcd123456’ into a bar code symbol you could use Set B, which would essentially use 11 bar code characters (one for the ‘Shift B’ and 10 for the characters) or you could use only nine bar code characters (one for the ‘Shift B’, four for the alpha characters, one for the ‘Shift C’ and three for the numeric characters)! The latter is obviously the better choice.

Don’t worry too much about all of this. Most good-quality bar code printing programs automatically determine which character sets to use for the greatest efficiency. Your bar code reader does the rest, so don’t lose any sleep over this one!

Advantages: Code 128 is the most space-efficient bar code in current existence which offers as large a character set.

Disadvantages: Some older or cheap bar code readers don’t read Code 128. Writing software to print Code 128 symbols can be a bit of a nightmare because of the three character sets.

Uses and Industries: Retail (for carton labeling), health care, industrial and commercial uses.

Interleaved 2 of 5 Code

Synopsis: This is a numeric-only bar code of variable length which is very space-efficient. Some faults with the symbology may affect its reliability but can be overcome with the use of Modulus Check Digit or length-checking of the symbol.

Technical Details: Interleaved 2 of 5 Code is rather interesting as it is an *interleaved* bar code. While most bar code symbologies read the widths of the bars, this code also reads the spaces between the bars. What this means is that the first character of the symbol is interpreted in the bars and the second in the spaces between those bars and so on through the other characters of the symbol. As you may have figured out, this means that every I 2/5 symbol must be an even number of characters in length. Often, a zero is automatically added to the beginning to make this work out.

Unfortunately, I 2/5 is not the most reliable symbol and it is possible to scan a partial symbol and have it interpreted as whole. While this is almost impossible with wand (pen) scanners, it can happen with laser scanning devices. To get around this flaw, you could set the bar code reader to only scan symbols of a given length, but that would restrict you to only scanning fixed-length bar codes. Alternately, a Modulus Check Character (MCC) could be added to the end of the symbol.

Advantages: Very compact and variable in length.

Disadvantages: Somewhat restrictive (i.e., numeric, even number of characters) and not very reliable unless length-checking or check digit calculations are made.

Uses and Industries: Retail case marking, printed circuit board manufacturing

UPC

Synopsis: Both a symbology and a standard for the retail industry, UPC has no place other than in the retail environment. A fixed-length, numeric-only code, it is quite space-efficient and is the granddaddy of bar code symbols.

Technical Details: UPC (Universal Product Code) has been an evolving symbology and standard, fraught with confusion and misunderstanding. So, if you really need to know, get ready to be confused!

Essentially, there are two basic types of UPC symbols to which another two "extensions" may be added.

The UPC-A symbol is the most common. It is a 12-digit symbol and is found on 90% of the grocery and retail products sold in North America. The first digit of the symbol designates the "number system". That is, if it is 0,7 or 8, it is designated as a fixed product (e.g., a 500g box of Kellogg's Rice Krispies). If the number system is a '1', the symbol designates a product of variable weight and/or size (e.g., 450g of lean ground beef). Finally, a '5' in the first position of the symbol designates that the symbol is on a coupon, which in bookkeeping terms, is the opposite of a product.

The second-through-sixth characters (five digits) of the UPC-A code designate the manufacturer of the product. All manufacturers wishing to use UPC codes on their products must apply for this number and pay $300 to a Canadian entity called the GPMC. The GPMC administrates the distribution manufacturer's numbers and see to it that no two companies have the same five-digit manufacturer's code anywhere in North America.

The characters in the seventh-through-eleventh positions designate the actual product code of the individual product. Therefore, these five digits would change from Kellogg's Rice Krispies to Kellogg's Corn Flakes, just as they would from the large box of Cheerios to the small box.

The last (12th) digit is what's known as a *Modulus 10 Check Digit* and is essentially the result of a rather drawn-out mathematical equation involving the first eleven digits. It was originally there to check the reliability of the bar code as it is scanned. Nowadays, with the reliability of the UPC symbol proven, it is really only there to check the undependable human who may have to key in the number manually, should the symbol be unscannable!

The UPC-E symbol was invented for those products and manufacturers who could not fit an entire UPC-A symbol on their package. Suffice it to say there were a limited number of UPC-E codes to go around and they were all gone long ago! UPC-E is a six-digit symbol, which is essentially a subset of the 12-character UPC-A with a bunch of zeros taken out. Therefore, a UPC-E symbol, when scanned, can be interpreted and transmitted as a six-digit product code or "expanded" into a 12-digit UPC-A code by automatically inserting the missing zeros.

Over the past few years the magazine and publishing industries have adopted the UPC code as their standard. Rather than leave well enough alone, they have begun to use "extensions" to the UPC-A and UPC-E symbols which add another 2- or 5-digit symbol off to the right side of the main symbol. Usually, this extension contains information such as the issue number of a periodical or the selling price of a mass-market paperback book.

Advantages: Small and efficient bar code symbol, excellently suited to its task in the retail industry.

Disadvantages: Very restrictive (i.e., fixed-length, many quality standards) and unworkable outside of the retail marketplace.

Uses and Industries: Retail grocery, book and magazine, garment, drug, etc.