Brain, M. (2015). How UPC bar codes work (First of four pages). Retrieved on 23 Oct 2015\* from <http://electronics.howstuffworks.com/gadgets/high-tech-gadgets/upc.htm>

How UPC Bar Codes Work

If you go look in your refrigerator or pantry right now, you will find that just about every package you see has a UPC bar code printed on it. In fact, nearly every item that you purchase from a grocery store, department store and mass merchandiser has a UPC bar code on it somewhere.

Have you ever wondered where these codes come from and what they mean? In this article, we will solve this mystery so that you can decode any UPC code you come across.

"UPC" stands for Universal Product Code. UPC bar codes were originally created to help grocery stores speed up the checkout process and keep better track of inventory, but the system quickly spread to all other retail products because it was so successful.

UPCs originate with a company called the Uniform Code Council (UCC). A manufacturer applies to the UCC for permission to enter the UPC system. The manufacturer pays an annual fee for the privilege. In return, the UCC issues the manufacturer a six-digit manufacturer identification number and provides guidelines on how to use it. You can see the manufacturer identification number in any standard 12-digit UPC code. The UPC symbol has two parts:

•The machine-readable bar code

•The human-readable 12-digit UPC number

The manufacturer identification number is the first six digits of the UPC number -- 639382 in the image above. The next five digits -- 00039 -- are the item number. A person employed by the manufacturer, called the UPC coordinator, is responsible for assigning item numbers to products, making sure the same code is not used on more than one product, retiring codes as products are removed from the product line, etc.

In general, every item the manufacturer sells, as well as every size package and every repackaging of the item, needs a different item code. So a 12-ounce can of Coke needs a different item number than a 16-ounce bottle of Coke, as does a 6-pack of 12-ounce cans, a 12-pack, a 24-can case, and so on. It is the job of the UPC coordinator to keep all of these numbers straight!

The last digit of the UPC code is called a check digit. This digit lets the scanner determine if it scanned the number correctly or not. Here is how the check digit is calculated for the other 11 digits, using the code 63938200039 from "The Teenager's Guide to the Real World" example shown above:

1. Add together the value of all of the digits in odd positions (digits 1, 3, 5, 7, 9 and 11). 6 + 9 + 8 + 0 + 0 + 9 = 32

2. Multiply that number by 3. 32 \* 3 = 96

3. Add together the value of all of the digits in even positions (digits 2, 4, 6, 8 and 10). 3 + 3 + 2 + 0 + 3 = 11

4. Add this sum to the value in step 2. 96 + 11 = 107

5. Take the number in Step 4. To create the check digit, determine the number that, when added to the number in step 4, is a multiple of 10. 107 + 3 = 110 The check digit is therefore 3.

Each time the scanner scans an item, it performs this calculation. If the check digit it calculates is different from the check digit it reads, the scanner knows that something went wrong and the item needs to be rescanned.

How is the Price Determined?

As you can see, there is no price information encoded in a bar code. When the scanner at the checkout line scans a product, the cash register sends the UPC number to the store's central **POS** (point of sale) computer to look up the UPC number. The central computer sends back the actual price of the item at that moment.

This approach allows the store to change the price whenever it wants, for example to reflect sale prices. If the price were encoded in the bar code, prices could never change. On the other hand, not encoding a fixed price gives the store an easy way to rip off customers. When you hear about "**scanner fraud**" in the news, that is what the newsperson is talking about. It is incredibly easy for a store to mistakenly or purposefully overprice an item.

One thing you will notice if you start looking at UPC codes in detail is that the big manufactures have manufacturer IDs with lots of zeros in them. Here are a few:

* **Post** - 043000
* **General Mills** - 016000
* **Del Monte** - 024000
* **Quaker Oats** - 030000

Coke's manufacturer ID is 049000 (pictured above). However, if you look at can of Coke or most 2-liter bottles, you will find that the UPC code is much shorter than on a 3-liter bottle -- only eight digits total. Here's the bar code from a 2-liter bottle of Sprite:



Zero-suppressed UPC code on a bottle of Sprite

©2000 HowStuffWorks

These short bar codes are called **zero-suppressed numbers**. There's a set of rules around forming zero-suppressed numbers from full numbers, but the basic idea is to leave out a set of four digits, all zeros. In the case of the Sprite UPC code, the 049 at the beginning is the first three digits of Coke's 049000 manufacturer ID. The 551 is the item number for this bottle of Sprite, shortened from 00551. The zero in the second-to-last digit is the fourth digit from Coke's manufacturer ID. The final digit is the normal check digit. The main reason for having zero-suppressed numbers is to create smaller bar codes for small product packages like 12-ounce cans.

The first digit of the manufacturer's identification number is special. It is called the **number system character**. The following data shows you what different number system characters mean:

* 0 = Standard UPC number (must have a zero to do zero-suppressed numbers)
* 1 = Reserved
* 2 = Random-weight items (fruits, vegetables, meats, etc.)
* 3 = Pharmaceuticals
* 4 = In-store marketing for retailers (A store can set up its own code, but no other store will understand them.)
* 5 = Coupons
* 6 = Standard UPC number
* 7 = Standard UPC number
* 8 = Reserved
* 9 = Reserved

Here is an example of a pharmaceutical bar code (number system character **3**), this one from a 4-ounce bottle of Selsun Blue dandruff shampoo:



**What is a Coupon Code?**

The **coupon code** is interesting (number system character **5**). If you have ever wondered how the scanner can read a coupon and reject it if you haven't bought the product, here's your explanation.



Coupon code for Post Honey Nut Shredded Wheat

©2000 HowStuffWorks

If you look at the coupon code, you can see that the coupon's bar code starts with a 5 to indicate that it is a coupon. The 43000 is Post's manufacturer ID. The next three digits (186) are called the **family code**. The next two digits (70) are a **value code**. The final digit is the normal check digit.

The family code and value code are set up arbitrarily by the UPC coordinator for the manufacturer. It must be done that way because a coupon will often be usable for a whole family of products. For example, a coupon might be good for four different kinds of soap made by the same manufacturer. In the same way, the value code represents the value of the coupon arbitrarily. The manufacturer sends the retailer the data that tells the retailer's computer exactly which products fit the family code, and exactly how much to take off. When the coupon is scanned, the POS computer:

1. Decodes the family code
2. Checks to make sure the customer purchased an item from the family
3. Decodes the value code
4. Sends the discount back to the cash register

**Can I Decode the Bars?**

So let's say you would like to decode the actual bars in the bar code and map them to numbers. This is something that will make you cross-eyed, but it can be done.

First of all, look at any 12-digit bar code. It is made up of black bars and white spaces between the bars. Assume that the thinnest bar or space that you see (for example, the first bar on the left) can be called "one unit wide." The bars and spaces can therefore be seen to have proportional widths of one, two, three or four units. If you look at any bar code you can see examples of these four widths.

The start of any bar code is "1-1-1." That is, starting at the left you find a one-unit-wide black bar followed by a one-unit-wide white space followed by a one-unit-wide black bar (bar-space-bar). Following the start code, the digits are encoded like this:

0 = 3-2-1-1

1 = 2-2-2-1

2 = 2-1-2-2

3 = 1-4-1-1

4 = 1-1-3-2

5 = 1-2-3-1

6 = 1-1-1-4

7 = 1-3-1-2

8 = 1-2-1-3

9 = 3-1-1-2

(Something to notice: All of these encodings seem to add up to 7.)

So let's take this barcode as an example:

The code embedded in the bars is **043000181706**:

* The bar code starts with the standard start code of 1-1-1 (bar-space-bar).
* The zero is 3-2-1-1 (space-bar-space-bar).
* The four is 1-1-3-2 (space-bar-space-bar).
* The three is 1-4-1-1 (space-bar-space-bar).
* The next three zeros are 3-2-1-1 (space-bar-space-bar).
* In the middle there is a standard 1-1-1-1-1 (space-bar-space-bar-space), which is important because it means the numbers on the right are optically inverted!
* The one is 2-2-2-1 (bar-space-bar-space).
* The eight is 1-2-1-3 (bar-space-bar-space).
* The one is 2-2-2-1 (bar-space-bar-space).
* The seven is 1-3-1-2 (bar-space-bar-space).
* The zero is 3-2-1-1 (bar-space-bar-space).
* The six is 1-1-1-4 (bar-space-bar-space).
* The stop character is a 1-1-1 (bar-space-bar).

Have fun decoding those 12-digit bar codes!