LZW compression, named after its three founders, Abraham Lempel, Jacob Ziv, and Terry Welch showcases a modern means of shrinking a file, despite being discovered shortly after the personal computer. Despite its rather long existence, it still impacts contemporary technology through the GIF, TIFF, and PDF file formats. Its simple and straightforward implementation has assisted in its success through the years.

First, LZW compression relies heavily on an underlying symbol table. The symbol table’s implementation is insignificant, but the quickest access and insert times possible are highly recommended. This table is initialized with all 256 (28 because of 8-bits) ASCII characters; these will be the initial “substrings” for the current uncompressed file. Then, a program loops through the file. For every iteration, the beginning of the file is scanned for the longest codeword contained within the codebook, or symbol table; this is called the longest prefix. So, for example, if the file began as “a little boy once was […]”, with the initialized symbol table, the longest prefix in this case would be “a”. Then, that prefix, concatenated with its neighbor character to the right, are placed into the symbol table with its appropriate index; in this case, that new codeword is “a “ with the codeword value 257 because a space is the next character in line. Depending on the language chosen to implement LZW, the file, or ‘input’ string, is shortened and moved past the encoded codeword (in this situation, “a”). Now, the input string reads “ little boy once was […]”. One critical step in this compression is actually writing out the compressed data to a file. Because the codeword found in this example is “a”, which happens to be the value 97 in ASCII (the initialized codebook), that value is written to a new file in binary with N-bits. In a fixed-length encoding, where N is the number of encoded bits, the codebook size never exceeds 2N, and every codeword written to the compressed file is N-bits. However, in a variable-length encoding, one may start with 8-bit codewords and move up to 16-bits, allowing for 216 (65,536) codewords.