Lexicographic Order and Linear Algebra

Sebastian R. Papanikolaou Costa

August 7, 2021

Abstract

In this short paper I will attempt to present a scheme that can be used order to an unordered set of length n binary strings lexicographically. We will use linear algebraic techniques to show that strings of 1's and 0's will indeed map to their natural number counterparts.

1 Order

In mathematics, the concept of an order can mean a lot of things. We have derivatives of increasing and decreasing orders in calculus as well as the ordering in a strictly monotonically increasing (or decreasing) sequence. We will concern ourselves with the latter.

Consider the set B_2 of binary strings of length 2:

$$B_2 = \{00, 01, 10, 11\}.$$

We know that by the technique of binary counting, each length-2 string corresponds to a natural number by considering the positions of the string to be powers of 2, from right to left, where a position with a 1 denotes the use of the power of 2 in that position and a 0 denotes the absence of the power of 2 in that position. To calculate the number that corresponds to the binary string, we simply add the powers of 2 that correspond to the 1's and 0's positions in the string, like so:

$$011010 \longrightarrow 0 + 2^4 + 2^3 + 0 + 2^1 + 0 = 26,$$
$$01 \longrightarrow 0 + 2^0 = 1,$$
$$0111 \longrightarrow 0 + 2^2 + 2^1 + 2^0 = 7.$$

For the base-2 system, this "ordering" is known as the *lexicographic* or *dictionary* order. We use simple matrix algebra to show that these strings will map to their natural number counterparts via some function, but first we must impose the constraint that all strings considered are of the same length.

2 The Order Encoding Scheme

Let's take a wild leap of faith here and take binary strings of length n to be vectors of a dimension n, mathematically we will need the vectors to be over the field of natural numbers $\mathbf{N}$ (or $\mathbf{Z}^+$). Furthermore, our vectors will contain only 1's and 0's.

Definition 1 (Order Encoding Function)**.** *Let* Υ *be a subspace of* $\mathbf{N}^n$*, or* $\Upsilon \subset N^n$*. Furthermore for all* $\upsilon_i \in \Upsilon$*, we have* $\upsilon_{i,j} = 0 \vee \upsilon_{i,j} = 1$ *for* $1 \leq j \leq n$ *and* $1 \leq i \leq |\Upsilon|$*. Then there exists a function* T *such that* $T : \Upsilon \longrightarrow \mathbf{N}$ *and we define it as follows:*

$$T = \left(\begin{bmatrix} 2^{n-1} & 0 & \dots \\ \vdots & \ddots & \\ 0 & & 2^0 \end{bmatrix} \begin{bmatrix} \upsilon_1 \\ \upsilon_2 \\ \vdots \\ \upsilon_n \end{bmatrix}_i \right) \bullet \begin{bmatrix} \upsilon_1 \\ \upsilon_2 \\ \vdots \\ \upsilon_n \end{bmatrix}_i \tag{1}$$

T is called the "Order Enoder" or "Order Enconding Function" and will map a binary string of length n to its correspong natural number representation.

Remark 1. *Note that we can "tighten" our restrictions on* Υ *by stating that* $|\upsilon_i|^2 \leq n$*.*

A brief example is given for sake of completion:

Example 1. *Consider* $\Upsilon = \{(0,0),(0,1),(1,0),(1,1)\}$*. Applying* T *an all elements of* Υ*, we get:*

$$\upsilon_1 = (0,0) \longrightarrow T = \left(\begin{bmatrix} 2^1 & 0 \\ 0 & 2^0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) \bullet \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0$$

$$\upsilon_2 = (0,1) \longrightarrow T = \left(\begin{bmatrix} 2^1 & 0 \\ 0 & 2^0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \bullet \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1$$

$$\upsilon_3 = (1,0) \longrightarrow T = \left(\begin{bmatrix} 2^1 & 0 \\ 0 & 2^0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \bullet \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2$$

$$\upsilon_4 = (1,1) \longrightarrow T = \left(\begin{bmatrix} 2^1 & 0 \\ 0 & 2^0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \bullet \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3$$

3 Conclusions

A proof need not be made but if necessary, one can show this using an induction process on n. The operations performed are well defined operations on vectors as well. On a side note, T is **not** invertible but is $1 - to - 1$, making it an injective mapping. T *may* be useful due to the fact that ordering binary strings could be handled with quick matrix algebra. However, it must be mentioned that this will work only for binary strings. The diagonal matrix in T can be studied independently as well, as we can generalize the matrix to hold powers of any prime number.

References

[1] Wikipedia contributors. Lexicographic order. Wikipedia, The Free Encyclopedia. July 24, 2021, 20:14 UTC. Available at: https://en.wikipedia.org/w/index.php?title=Lexicographic_order&oldid=1035290692. Accessed August 7, 2021.