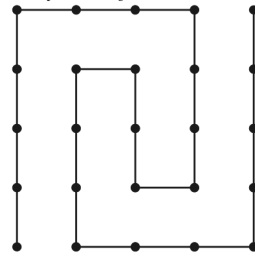

SOLUTIONS

A Twisted Enumeration of the Positive Integers

11733 [2013, 854]. *Proposed by Donald Knuth, Stanford University, Stanford, CA.*

Let $V = \{0, 1, 2, 3, 4\}^2$. Say that nonnegative integers a and b are *adjacent* when their base-5 expansions $\cdots a_2 a_1 a_0$ and $\cdots b_2 b_1 b_0$ satisfy the condition that if $i > j \geq 0$ and $(a_i, a_j) \neq (b_i, b_j)$, then (a_i, a_j) and (b_i, b_j) are consecutive in the path through V shown at right (horizontal coordinate listed first). Thus, for example, 0 is adjacent to 1. Similarly 48 (expansion 143_5) is adjacent to 47 (expansion 142_5) and 73 (expansion 243_5).



- (a) Prove that every positive integer is adjacent to exactly two nonnegative integers.
- (b) Prove that with this definition of adjacency, the nonnegative integers form a path $\langle x_0, x_1, x_2, \dots \rangle$ starting with $x_0 = 0$.
- (c) Explain how to compute efficiently from n the number x_n that comes n steps after 0, and determine $x_{1,000,000}$.

Solution by Richard Stong, Center for Communications Research, San Diego, CA. We first define a *raster path* P_d from 1^d to 3^d in $\{1, 2, 3\}^d$. For $d = 0$ the path is the single vertex given by a 0-tuple. Denote by aP the path given by prepending a to every vertex of a path P . For $d \geq 1$, the path P_d is the concatenation of $1P_{d-1}$, $2P'_{d-1}$, and $3P_{d-1}$, where P' denotes the reverse of P . By induction on d , every step along the path moves one digit up or down by 1. Call a step an $i|j$ step if it changes i to j or j to i in one position.

To find the n th vertex of P_d (starting from 0), write n in base 3, replace the k th digit by its difference from 2 if the number of 1s to the left is odd, and then increment each digit by 1. When two vertices differ only in position k , having 1 and 2 there, they are adjacent via a 1|2 step in P_d if and only if all entries to the right are 3. When in position k they have 2 and 3, they are adjacent via a 3|2 step in P_d if and only if all entries to the right are 1. Because P_d has 1^d at the beginning and 3^d at the end, it follows by induction on d that 1|2 steps alternate with 3|2 steps.

We next define a *double raster path* D_d from 0^d to 4^d in $\{0, 1, 2, 3, 4\}^d$. Begin with the path in $\{0, m, 4\}^d$ obtained from P_d by replacing every 1 with 0, every 2 with m , and every 3 with 4. Steps become $0|m$ steps or $4|m$ steps, alternating. Turn D_d into a path through $\{0, 1, 2, 3, 4\}^d$ as follows: When entering an element v of $\{0, m, 4\}^d$ with j entries equal to m , expand v into a copy of P_j in the coordinates equal to m , leaving the other coordinates fixed. If we enter v by a $0|m$ step, then we follow P_j from 1^j to 3^j . If we enter v by a $4|m$ step, then we follow P_j' from 3^j to 1^j . For example, the state $0m04m$ follows $0m044$ and expands to $03043, 03042, 03041, 02041, 02042, 02043, 01043, 01042, 01041$.

We summarize important properties of double raster paths, all of which can be verified by induction on d . Exactly two of properties (3)–(6) apply at each vertex other than 0^d and 4^d (where exactly one applies), governing the steps entering and leaving that vertex. Left and right are relative to the coordinate that changes in a step.

- (1) All steps change a single coordinate to an adjacent value.
- (2) The raster path is followed forward in the blocks arising from elements whose number of copies of m is odd and backwards in those where it is even.
- (3) A vertex is involved in a $0|1$ step when each entry to the right is 4 and each entry to the left is 0, 1, or 4.
- (4) A vertex is involved in a $1|2$ step when each entry to the right is 0, 3, or 4.
- (5) A vertex is involved in a $2|3$ step when each entry to the right is 0, 1, or 4.
- (6) A vertex is involved in a $3|4$ step when each entry to the right is 0 and each entry to the left is 0, 3, or 4.

For example, when $d = 4$ the vertex 2233 with leftmost entry 2 cannot satisfy (3) or (6), so its transitions must involve 2; its neighbors are 2133 and 2232.

A position cannot first change away from 0 until each entry to the right is 4. Thus the first 5^{d-1} steps of the double raster path D_d are a copy of D_{d-1} through vertices with leading coordinate 0. Hence the double raster paths of all dimensions combine to give a single path D_∞ through $\{0, 1, 2, 3, 4\}^\infty$.

We now return to the given problem. Each edge in the graph shown corresponds to changing the entry in one position to a neighboring value. An edge of V that changes a vertical coordinate from 0 to 1 has horizontal coordinate 0, 1, or 4. An edge of V that changes a horizontal coordinate from 0 to 1 has vertical coordinate 4. Thus two integers that differ in a single digit where one has a 0 and the other a 1 are adjacent if and only if all digits to the left are 0, 1, or 4 and all digits to the right are 4. Comparing this to (3), we see that they are adjacent if and only if they are consecutive in D_∞ . Analyzing the other three cases similarly and comparing them to (4), (5), and (6), we see that nonnegative integers are adjacent if and only if they are consecutive in D_∞ . This establishes parts (a) and (b) of the problem.

For part (c), let y_k be the string of symbols from $\{0, m, 4\}$ corresponding to the vertex x_k . If k has d digits in base 5, then y_k will have d characters (since D_∞ starts with D_d). If the leading digit of k is 4, say $k = 4 \cdot 5^{d-1} + k'$, then y_k will be given by prepending a 4 to $y_{k'}$ (with leading zeroes added if necessary to make $y_{k'}$ of length $d - 1$), since the last segment of D_d is $4D_{d-1}$. If the leading digit is 1, 2, or 3, then let

$s = \lfloor (k - 5^d)/3 \rfloor < 5^{d-1}$. Now y_k will be given by prepending m to the complement (swapping symbols 0 and 4) of y_s , since the middle section of D_d is mD'_{d-1} and every block has three times as many vertices as the corresponding block of D_{d-1} due to the extra m . Running this calculation backwards, we can find the least j such that $y_j = y_k$. Hence we know how far into the raster path k is. Specifically, take the string y_k and swap the symbols 0 and 4 whenever there is an odd number of symbols m to the left. Call the resulting sequence $z_{d-1} \cdots z_0$. Define $k_0 = 0$ and

$$k_{j+1} = \begin{cases} k_j & \text{if } z_j = 0; \\ 3k_j + 5^j & \text{if } z_j = m; \\ k_j + 4 \cdot 5^j & \text{if } z_j = 4. \end{cases}$$

Now k_d is the least term in the block. From (2) we know whether the raster path is followed forwards or backwards, so we can compute the values for these coordinates as described in the opening paragraph and use them to find x_k .

For the specific case $k = 1000000$, we compute as follows. Letting an overbar denote complementation, and m^k to denote k consecutive copies of m , we obtain

$$\begin{aligned} y_{1000000} &= m\overline{y_{203125}} = m^2 y_{41666} = m^3 \overline{y_{8680}} = m^4 y_{1851} = m^5 \overline{y_{408}} \\ &= m^6 y_{94} = m^7 \overline{y_{23}} = m^7 0\overline{y_3} = m^7 0m. \end{aligned}$$

We then calculate $k_1 = 1$, $k_2 = 21$, $k_3 = 88$, $k_4 = 389$, $k_5 = 1792$, $k_6 = 8501$, $k_7 = 41128$, $k_8 = 201509$, and $k_9 = 995152$. Thus, the first step in this block occurs at index 995152, and $x_{1000000}$ is the 4848th term in this block. We follow the raster path backwards in this block, because it has eight symbols m . Writing $4848 = 20122120_3$ in base 3 and complementing every digit with an odd number of 1s to its left gives 20100120 . Hence the 4848th step of the raster path P_8 is 31211231 and the 4848th step counting backwards from the end is the complement 13233213 . Hence $x_{1000000} = 132332103_5 = 667778$.

Also solved by Y. J. Ionin, O. P. Lossers (Netherlands), and the proposer. Parts (a) and (b) also solved by TCDmath Problem Group (Ireland).