ERRATA CORRIGE

"Pearls of Algorithm Engineering"

Paolo Ferragina, Cambridge University Press, 2023

https://www.cambridge.org/core/books/pearls-of-algorithm-engineering/95061352D7263CCCBD4F243018236EB2

CHAPTER 3

- Page 31, reference [3]: the term "Complexity" should be lowercase

CHAPTER 6

- Page 78, the definition of Δ_j should be just 2^{k-1} that is surely at most n.

CHAPTER 7

- Page 84, line +5 of section 7.2.1: Change "Since buckets 1 and 6" to "Since buckets 1 and 7"
- Page 85, last line: substitute 666 with 766.

CHAPTER 8

- Page 103, proof of Theorem 8.3: Q(k) is the probability that **at least** k keys are hashed to a specific slot (as actually this is the use made of Q(k) in the rest of the proof). Moreover, few lines below it is introduced a random variable X_j that is never used in the rest of the proof, but it was intended as Q(k) = P(X_j >= k), for a given slot j.
- Page 105, line -6: the inverse modulo *m* of $(x_0 - y_0)$ is in [1, m-1]
- Page 106, line +11: the condition on *a* can be expressed better as >0, as then indicated in the definition of $H_{p,m}$ in the line below.

CHAPTER 9

- Page 143, line +7 from paragraph starting with *"A running example..."*. The occurrence of s_8 should be s_7.
- Proof of Theorem 9.5, line +2, the phrase "no more than b/g" should be "no more than \lceil b/g \rceil".
- Proof of Theorem 9.5, pag. 138, when upper bounding |B_i|, change "we can write" into "we can write (by assuming that g divides b, for the sake of presentation):"

CHAPTER 11

- Page 203, just line -5 before the beginning of Section 11.5: In the phrase starting with "In particular, if this distribution is concentrated…", the two occurrences of *s* should be substituted with *c*.

- Page 204 in Section 11.5 (middle of the page): it is written that "[..] we have m-l distinct elements of S' and the smallest one is larger than low [..]". To be clearer, it should be written "larger or equal than" as reported in the following formula.

- Page 206 in Section 11.6 (middle of the page): it is written that "Elias-Fano code works on monotonically increasing sequences". It should be written "monotonically nondecreasing sequences" and thus manage sets with repeated items. However, I mention that in the context of inverted lists for search engines, the case is the one of monotonically increasing sequences.

- Page 206 in Section 11.6 (bottom of the page): There is a counting on the number of 0s in the binary sequence H, which states that they are bounded by n. So the next phrase about the bit length of H should state that it is "bounded by" 2n, and not that it is exactly 2n. So the number of 0s is <u>at most</u> the number of 1s.
  Actually, to make this argument hold for every value of maxS and n, we should represent only the configurations up to maxS (included), and encode the length of H into an integer, which is stored anyway by the system to correctly access H. In fact, if n=15 and maxS = 19, we then have \ell = \ceil \log_2 (u/n) \ceil = \ceil \log_2 1,.... \ceil = 1, and h = 4. Thus the overall number of configurations over h bits is 2^h = 16 (= n+1 and thus larger than n).

  But if we limit ourselves to represent only the configurations up to maxS, the maximum bucket value is actually (\floor maxS/2^\ell \floor), and thus the number of buckets is (\floor maxS/2^\ell \floor + 1) because we have to include the configuration 0^h. This formula then expresses the maximum number of 0s in H, and it is bounded above by n as follows:

  (\floor maxS/2^\ell \floor + 1) <= (\floor (u-1)/2^\ell \floor + 1)
  <= (\floor (u-1)/2^{\ceil log_2 (u/n) \ceil} \floor + 1)
  <= (\floor (u-1)/2^{log_2 (u/n)} \floor + 1)
  <= (\floor (u-1)/(u/n) \floor + 1) = (\floor n - (1/(u/n)) \floor + 1) = (n-1) + 1 = n

  Recalling the previous example with n=15 and maxS = 19 = $(10011)_2$, we then have that the largest configuration is the H-part of maxS, hence 9 = $(1001)_2$. Thus the number of 0s is 9+1 = 10 < 15 = n; and, in fact, the formula above returns:
  (\floor maxS/2^\ell \floor + 1) = (\floor 19/2 \floor + 1) = 10.
  As another example (which gets exactly n configurations of 0s), take the one in Figure 11.8. So n=8, maxS = 31 = $(11111)_2$, and thus u = 31+1 = 32, \ell = 2, and h=3. Thus the largest high part is $(111)_2$ which means 8 = 2^3 buckets = n; and in fact the formula above returns:
  (\floor maxS/2^\ell \floor + 1) = \floor 31/4 \floor + 1 = 8.

CHAPTER 15

- Page 278, line -1, and Page 279 line +1: both occurrences of B[26] should be B[17].
- Page 281, line +5 in step 6, the modulo function should be on k (i.e., number of 1s in a small block) and not on k^2.
- Page 288, line +5: Rank1(13) should be 8 instead of 6.