The Self-Referential Sentence

Question

In chapter 2 of his book "Tracking the Automatic Ant", David Gale reports that Lee Sallows had found this sentence:

> This computer-generated sentence contains two hundred forty seven letters: four *a*'s, one *b*, four *c*'s, five *d'*s, forty four *e*'s, nine *f'*s, three *g*'s, seven *h*'s, eleven *i*'s, one *j*, one *k*, three *l'*s, two *m*'s, twenty nine *n*'s, nineteen *o*'s, two *p*'s, one *q*, fourteen *r*'s, thirty one *s*'s, twenty five *t'*s, seven *u*'s, eight *v*'s, seven *w*'s, two *x*'s, six *y*'s, and one *z*.

You will notice that it indeed correctly describes itself.

Is this the only such sentence? Or are there any others?

Answer

The new one is shorter. Here it is:

> This computer-generated sentence contains two hundred forty three letters: four *a*'s, one *b*, four *c*'s, five *d*'s, thirty seven *e*'s, ten *f'*s, five *g*'s, nine *h*'s, fourteen *i'*s, one *j*, one *k*, two *l'*s, two *m*'s, twenty four *n*'s, nineteen *o*'s, two *p*'s, one *q*, fourteen *r*'s, twenty eight *s*'s, thirty *t*'s, eight *u*'s, five *v*'s, eight *w*'s, two *x*'s, six *y*'s, and one *z*.

I used the conventions of the original sentence, in that there is no word "and" after "hundred".

In the book, there is a brief description of how it was found, which is by adjusting one of the numbers at random to its possibly correct value, and re-evaluating. If we keep trying randomly selected elements of the text until it eventually makes correct sense, he claims it will eventually settle to a solution. I wondered how many tries might be necessary to try, and set out to write a program to do exactly what he suggests.

It took a few tens of millions of spins until it settled, but what surprised me was that it had found a sentence different from Lee Sallows'. On later tries with different seeds for the random number generator, there were a preponderance of the new one over Sallows', and the number of adjustments that had to be made varied from just over 1,000,000 to more than 100,000,000 in the tries I used.

Are there any more? What about other languages?

Over to you to investigate.

Addendum: Since writing the above, I found that the name for these types of sentences is *autogram*, and this website gives many more examples in various languages: http://autograms.net/

I also tried Esperanto as the language, but the first attempt never settled, even after a billion spins of the algorithm. However, changing the introductory part of the sentence, did provide a solution. The Esperanto alphabet consists of 28 letters; it does not contain *q, w, x* or *y,* but does have six additional letters with diacritics, *ĉ, ĝ, ĥ, ĵ, ŝ,* and *ŭ*.

Here it is:

Komputilo kreis ĉi tiun frazon kiu havas ducent kvindek tri literojn: dek unu *a*-oj, unu *b*, du *c*-oj, du *ĉ*-oj, dudek kvar *d*-oj, dek sep *e*-oj, du *f*-oj, unu *g*, unu *ĝ*, du *h*-oj, unu *ĥ*, dek kvar *i*-oj, dudek kvin *j*-oj, unu *ĵ*, dudek kvar *k*-oj, tri *l*-oj, du *m*-oj, dek sep *n*-oj, dudek ok *o*-oj, kvar *p*-oj, dek du *r*-oj, kvin *s*-oj, unu *ŝ*, naŭ *t*-oj, tridek *u*-oj, tri *ŭ*-oj, naŭ *v*-oj, kaj du *z*-oj.

Andy Pepperdine

2021-09-26