Same but different: “common-substring” vs “Kolmogorov complexity”

written by smooch on Functor Network
original link: https://functor.network/user/3157/entry/1256

I notice something of an inverse relationship between *“the common-substring problem’s solution”*, and *“Kolmogorov complexity”* — in the sense that *“each is comprised of (and defined by) the set of constituent elements, which are excluded and ignored by the other”*…

… but i can’t find this relationship discussed anywhere: *—might anyone point me towards the correct search-term, or some ‘accessible’ texts, on this relationship?*

To explain:

1. The solution to the *“common-substring problem”*, is all of the *“common words, and strings thereof”*, between two input texts

   *the “common-substring problem” is typically discussed in terms of finding only the “**longest** common-substring” — but rather here, we will be interested in “all commonality”*

2. *“Kolmogorov complexity”* frames the *“randomness”* and *“incompressibility”* of some sequence; and is based upon a process which *“replaces”* sub-strings which exist elsewhere *(typically framed as the result of some generator function)*, referred to as compression; leaving only the unrecognised elements behind, which are considered *“random” (and incompressible)*

So, if we consider our two *(input)* strings from the *“common-substring problem”*, imagine:

i. the first, as *“the source of candidate sub-strings for kolmogorov compression replacement”*, and

ii. the second, as *“the active sequence which is compressed by every candidate substring from (i) found in itself” — (leaving only elements which are not replaced)*

What I’m pointing at then, is the way in which:

a) for sequences a, b , with *“common”* elements c *(found $a \cap b = c$)*

b) the *"uncommon"* elements u , *(found by* $a \setminus c = u$ *, or* $b \setminus c = u$ *)*, equate to *"incompressible"* elements; which when measured, result in *"kolmogorov complexity"* k , where $|u| = k$

—*is this "correct"?*

Might anyone point me towards the correct search-term, or some accessible texts, on this relationship?

Many thanks!