On Geometrical Characteristics of String Similarity Metrics

Last updated: 24 December 2020

different types of edit distance
it's wikipedia page has very very nice things

Levenshtein Distance

definition
how do we visualize it
distribution of edit distance between two random strings
how does levenstein space look like
difference between rimennian geometry and euclidean
all of them can be formed into a graph problem using a matrix where the cost of non existent operations is infinity
shortest path in a dag problem
a matrix where it's mid-lines are
all of them map to finding the shortest path in a matrix
k-bounded edit distance

Hamming Distance

how do you see it?
some form of visualization
hamming distance cube

Ulam Distance

Applications in DNA Data Storage

penalizing it as a cost metric
when you try to minimize some metric
when your algorithm tries to minimize some metric, if you can assume that it's error is bounded, you can also assume that the distribution of it's error is on average something good
their applications to dna storage and especially trace reconstruction

References