Linear Strategy Distance: Measuring Strategy Similarity in Evolutionary Simple Dynamics

Walter Stover

George Mason University

03/07/2025

Keywords: Evolutionary Game Theory, Bounded Rationality and Learning, Outlier Analysis

Normalized compression distance (NCD) is a simple means of constructing a similarity index between two objects by measuring how difficult it is to transform one object into another, and has been used in a wide variety of clustering and classification applications, including anomaly detection, clustering heterogeneous data, network dynamics, and even music classification. The NCD itself is based on a more fundamental expression,

Where is the shortest computer program that transforms two strings into each other, and is the algorithmic information measure of given input . This is the *information distance* between these two strings. To normalize this, the *normalized information distance* (NID) is measured instead:

In practice, the NID is not fully computable. However, Vitanyi and Cilibrasi (2005) rewrites NID to rely on a compressor to construct the NCD as an approximation of the NID:

The better the compressor , the closer the NCD approximates the NID, and NCD takes on a value in a range between 0 and 1, where the higher the NCD score, the more dissimilar two strings are to each other, i.e. the more transformations are needed to turn one string into another. While there are other simpler distance measures such as Hamming distance that are fully computable and are especially compatible with binary representations, using the NCD allows us to capture more nuance in our similarity measure by including more detail at a later stage of this work. For instance, we can extend our similarity measure to also take into consideration the response of these strategies to complex histories. This will be useful, as we plan to extend Lindgren to allow for more complex strategies by relaxing memory length constraints,

We can apply NCD to measure the similarity between strategies played in evolutionary space, where strategies are encoded in simple binary form as in Lindgren (1991):

Where is the action taken by the player in the event of history . Given two strategies, apply a compressor such that:

Where and are the compressed forms of genetic "code" of each strategy.

Then we can apply the NCD definition above which will return a continuous [0,1] measurement of how similar these strategies are to each other. Because NCD is designed to run on longer strings, we repeat the string representation of these strategies multiple times before running it through this NCD measure.

If we run this NCD measure on the encoded representations of three basic strategies of always cooperate (ALLC), always defect (ALLD), and Tit-for-Tat (TFT), we can generate a strategy topography:

A graph of different colored bars

AI-generated content may be incorrect.

Fig. 1: Strategy Topography of ALLC, ALLD, TFT – each Strategy Index represents a strategy pairing.