Sparsity refers to the number of changes between the factual and counterfactual sequence. We typically want to minimize the number of changes. However, sparsity is hard to measure, as we cannot easily count the changes. There are two reasons why this is the case: First, the compared sequences can have varying lengths. Second, even if they were the same length, the events might not line up, so we can simply count the changes to a feature. Hence, we use the previously established Semi-strucured Damerau-Levenshtein distance (SSDLD) to solve this issue. The sparsity distance uses a cost function as specified in Equation 1.

$$cost(a_i, b_j) = \sum_d \mathbb{I}(a_{id} = b_{jd}) \tag{1}$$
$$a_i, b_j \in \mathbb{R}^d$$

Here, $\sum_d \mathbb{I}(a_{id} = b_{jd})$ is an indicator function, that is used to count the number of changes in a vector.