On depth functions and robustness in formal concept analysis

In this paper we introduce the notion of a depth function for the case where the data points are objects of a formal context. The underlying space is here the closure system of all formal concept extents. Since this space has no metric structure, known properties of a classical statistical depth function like e.g., *affine equivariance*, *maximality at the center*, *quasi-concavity*, *vanishing at infinity*, have to be redefined or modified for this non-metric situation, wherever possible. We analyse these modified properties for the generalization of Tukey's depth and the convex hull peeling depth. It shows up that Tukey's depth has most of the desirable properties. Furthermore, the only depth function which satisfies one additional property of reflecting a certain notion of betweenness is the peeling depth.

Beyond the analysis of structural properties we also analyze the breakdown point for Tukey's depth and the peeling depth. Compared to the case of a space like $\mathbb{R}^d$, which is naturally equipped with the Euclidean metric, in our non-metric situation we have to modify the classical notions of robustness, which is still possible. While in the classical case $\mathbb{R}^d$ the breakdown point of Tukey's depth is at least 1/(d+1), here we find that the breakdown point for both Tukey's depth and the peeling depth is closely related to the VC dimension h of the underlying closure system. Concretely, the breakdown point is always at least $1/h$ and this bound is sharp in a certain sense.

This insight also allows a form of regularization of depth functions in the sense of being able to enlarge the breakdown point by reducing the VC dimension. One special method is what we call double peeling depth. There, one does not only peel objects. In every step of object-peeling, one additionally also locally peels some large contranominal attribute-substructures that lead to a high VC dimesnion and that therefore mask other hidden substructures. An analysis of both Tukey's depth as well as the peeling depth for different concrete data sets that demand different kind of conceptual scaling will conclude the paper.