Do Euler diagrams approximate Card Sorting?

Stefan Schneider[1][000−0001−6572−5983] and
Andreas Nürnberger[0000−0003−4311−0624]

Data and Knowledge Engineering Group,
Faculty of Computer Science,
Otto von Guericke University Magdeburg,
Universitätsplatz 2, 39106 Magdeburg, Germany
stefan.schneider@sschneider.de, andreas.nuernberger@ovgu.de
http://www.dke.ovgu.de

Abstract. The Spatial Arrangement Method (SpAM) has been proposed as an efficient alternative to the Pairwise Rating Method (PRaM) for similarity space extraction of perceptual items (i.e. images). Recently, SpAM has been criticized as approximating PRaM, particularly for conceptual items for which high-dimensional data are typically available. Since PRaM and SpAM are typically evaluated at the instance level, the classification in the similarity space is typically based on a continuous measurement level. In contrast, at the concept level, it is convenient to create a discrete similarity space with corresponding alternative elicitation methods. In this work, we address the question of whether a visual elicitation method can be a reasonable approach to a non-visual elicitation method for conceptual items in the context of domain exploration. We evaluate the ability to use Card Sorting and Euler diagrams to retrieve relevant documents using a search engine.

Keywords: Card Sorting · Euler diagram · similarity space · Domain exploration.

How can similarity data be collected from human judgments? Traditionally, the most direct elicitation method for determining similarity judgments has been based on step-by-step similarity ratings [5, 11]. Here, a similarity rating can be made using a Likert scale [11] in which the subject must assign a numerical value expressing the degree of similarity (e.g., 1 for very similar and 9 for very different [5]) between a presented pair of items, the so-called Pairwise Rating Method (PRaM). The PRaM elicitation can only be applied to a very small context in relation to all (typically 2 of n or 3 of n) available objects. A great many repetitions are required to extract the entire similarity space of the elements. Extracting the similarity space of n elements requires $n(n-1)/2$ similarity judgments [4].

A more efficient method was presented by Goldstone, which organizes the similarities of objects based on their spatial arrangement - the so-called Spatial

Arrangement Method (SpAM) [4]. Participants are presented with a random arrangement of objects on a computer screen and arrange them in an appropriate manner, with the distances between objects reflecting the given degree of dissimilarity - similar objects are close together, while dissimilar objects are far apart. While a geometric space of objects and psychological distances can be generated as output using MultiDimensional Scaling (MDS) [10] on the basis of PRaM ratings, SpAM allows participants to create a geometric representation (Euclidean space) themselves that best matches their intuitions about item similarities. In evaluating the organization of several font examples, Goldstone found that SpAM correlates with PRaM [4]. Therefore, SpAM has been proposed as an interesting alternative to existing serial approaches that does not require repetition to maintain a user's similarity space. In recent years, SpAM has been popularly used for real-world use cases such as organizing movie or music collections [6, 9]. Despite the clear efficiency benefit of using SpAM instead of PRaM, the quality of the approximation has been questioned by Verheyen et. (2016) [11], who reanalysed the evaluation of Hout et. al. (2013) [5]. In Summary, Verheyen and colleagues recommend to use PRaM instead of SpAM, especially in cases where the similarity space consists of conceptual items containing high dimensional data with latent features that are not given for all participants.

Similarity spaces based on PRaM and SpAM assume continuous dimensions. It has already been pointed out that there are other survey methods that require only a discrete scale level [3, 11]. Our basic premise is that when structuring conceptual items (i.e., terms), a discrete scale level is sufficient to express differences in a given dimensional space. PRaM and SpAM are two examples of a non-visual and a visual elicitation method expressing a continuous similarity space. To create a discrete similarity space, we intend to compare Card Sorting (non-visual) and Euler diagram (visual) as two other examples of elicitation methods. In Card Sorting, without reference to specific criteria, participants are asked to sort cards - in our case conceptual elements (i.e., terms) - into piles so that elements in one pile are more similar to each other than elements in separate piles [12]. Euler diagrams are a popular visual representation that use curves to represent sets [2]. After adding a set on the screen, a participant can add elements to the set (i.e. terms) and label the set [7]. Euler diagrams can be used to build hierarchical structures that can be directly converted into concept lattices [7]. Repeated application of Card Sorting (also called hierarchical sorting [8]) also allows the construction of hierarchical structures, where attributes can be added in a further step, based on the question "In what way are these elements the alike?" [12]

We evaluate the impact of a user-defined similarity space on document ranking in the case of domain exploration, where almost all relevant documents of a given task must be identified by reading as few documents as possible. By representing the similarity spaces as concept lattices [3], we are able to determine a weighting factor that adapts the ranking of citations in the documents [1]. Starting from a reference publication (e.g., Ganter, Stumme (2004) [3]) with 200 documents (with the highest citation frequency) and 3 index terms per document, each participant has to extract contextually relevant terms from a given task context (e.g., "Please identify publications that primarily deal with formal concept analysis."). The contextually relevant terms selected by the participant must then be structured based on (a) hierarchical card sorting or (b) Euler diagrams. The representation of the elicitation methods is converted into concept lattices, which are used to calculate the term weights. These term weights are aggregated into document weights that adapt the ranking of the documents. Based on the assessment topic experts separately determined which documents are relevant for the given context.

We will present research results based on the degree of relevance of documents in ordered sequence by comparing Euler diagrams and hierarchical card sorting.

References

1. Beel, J., & Gipp, B. (2009, April). Google Scholar's ranking algorithm: the impact of citation counts (an empirical study). In 2009 third international conference on research challenges in information science (pp. 439-446). IEEE.
2. Chapman, P., Stapleton, G., Rodgers, P., Micallef, L., & Blake, A. (2014, July). Visualizing sets: An empirical comparison of diagram types. In International Conference on Theory and Application of Diagrams (pp. 146-160). Springer, Berlin, Heidelberg.
3. Ganter, B., Stumme, G., & Wille, R. (Eds.). (2005). Formal concept analysis: foundations and applications (Vol. 3626). springer.
4. Goldstone, R. (1994). An efficient method for obtaining similarity data. Behavior Research Methods, Instruments, & Computers, 26(4), 381-386.
5. Hout, M. C., Goldinger, S. D., & Ferguson, R. W. (2013). The versatility of SpAM: a fast, efficient, spatial method of data collection for multidimensional scaling. Journal of Experimental Psychology: General, 142(1), 256.
6. Low, T., Hentschel, C., Stober, S., Sack, H., & Nürnberger, A. (2017, January). Exploring large movie collections: Comparing visual berrypicking and traditional browsing. In International Conference on Multimedia Modeling (pp. 198-208). Springer, Cham.
7. Priss, U. (2020, September). Set Visualisations with Euler and Hasse Diagrams. In International Workshop on Graph Structures for Knowledge Representation and Reasoning (pp. 72-83). Springer, Cham.
8. Rugg, G., & McGeorge, P. (1997). The sorting techniques: a tutorial paper on card sorts, picture sorts and item sorts. Expert Systems, 14(2), 80-93.
9. Stober, S., Low, T., Gossen, T., & Nürnberger, A. (2013). Incremental Visualization of Growing Music Collections. In ISMIR (pp. 433-438).
10. Torgerson, W. S. (1965). Multidimensional scaling of similarity. Psychometrika, 30(4), 379-393.
11. Verheyen, S., Voorspoels, W., Vanpaemel, W., & Storms, G. (2016). Caveats for the spatial arrangement method: Comment on Hout, Goldinger, and Ferguson (2013).
12. Weller, S.C. and Romney, A.K. (1988). Systematic Data Collection. Sage Publications, Newbury Park, CA.