Towards Data-driven Conceptual Spaces

Hadi Banaee, Erik Schaffernicht, and Amy Loutfi

Centre for Applied Autonomous Sensor Systems, School of Science and Technology,
Örebro University, SE-701 82 Örebro, Sweden
{hadi.banaee,erik.schaffernicht,amy.loutfi}@oru.se

Motivation One goal of artificial intelligence is to construct systems that can model cognitive activities, such as concept learning and semantic inference [1]. However, a critical issue is how numerical information is to be modelled in knowledge representation frameworks [2]. Deriving descriptions for concepts is non-trivial in artificial systems especially if available information is provided in numeric or non-symbolic representations (e.g., sensor data).

The theory of *conceptual spaces* introduced by Gärdenfors [3] describes a mid-level representation to addressing both concept learning and semantic inference [1], and provides a possible solution to deriving descriptions of concepts. A conceptual space consists of *quality dimensions* in various *domains* that are placed within a geometrical structure to represent the *concepts* [3]. Conceptual spaces are principally derived in a *knowledge-driven* manner, based on the assumption that the prior knowledge from perceptual mechanisms or experts manually initialises the elements of the space (i.e., domains, quality dimensions, and concepts' regions) [4]. However, a question is whether the construction of a conceptual space can be automated, at least in part, in a data-driven manner [5] given a series of numerical measurements. Answering this question is strongly motivated by a growing class of problems that involves complicated observations and have little or no prior knowledge concerning their semantic significance [6].

This study considers the task of creating semantic representations to describe numerical observations (See Fig. 1). Our claim is that the data-driven conceptual spaces can be considered as a semantic representations to conceptualise the perceived information and to be utilised to infer linguistic descriptions.

Contributions The main contribution of this study is to investigate the possibility of inferring human understandable semantics for any given data set through data-driven conceptual spaces. Our proposed framework provides a procedure to utilise machine learning algorithms for the task of identifying relevant features and concepts in a numerical data set, in order to specify the domains and quality dimensions. The semantic inference is then introduced to linguistically represent a new observation within the built conceptual space. The framework is assessed by applying the proposed approaches on various data sets. The formal definitions and the proposed algorithms of the framework are elaborated in detail in [7].

Challenges The origin of quality dimensions is still an open question [3]. Once the process of constructing a conceptual space starts, as Quine noted [8], some in-

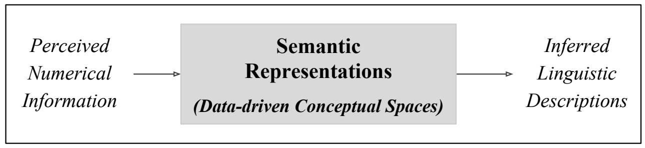

Fig. 1: Semantic representation for describing numerical data.

nate quality dimensions are needed to make *concept learning* possible. However, there is no unique way to specify which dimensions are sufficient to characterise the concepts. Another issue of constructing conceptual spaces in a data-driven way is the semantics of domains. There still exists the problem of verifying the semantic dependency of the quality dimensions to each other within a domain, analytically. Indeed, solving the issue of domain specification can lead to forming a general solution to the problem of determining an evaluation criterion to choose between competing conceptual spaces, an issue raised by Gärdenfors in [9].

The assumption of using machine learning techniques to construct conceptual spaces is that the more distinctive a feature is for a class in a numerical data set, the more descriptive its corresponding quality dimension is for the target concept within the created conceptual space. Thus, one possible future work can be the use of the proposed approaches to measuring the goodness of conceptual spaces, statistically.

References

1. Aisbett, J., Gibbon, G.: A general formulation of conceptual spaces as a meso level representation. Artificial Intelligence **133**(1) (2001) 189–232
2. Gärdenfors, P.: Conceptual spaces as a framework for knowledge representation. Mind and Matter **2**(2) (2004) 9–27
3. Gärdenfors, P.: Conceptual spaces: The geometry of thought. MIT press (2000)
4. Rickard, J.T., Aisbett, J., Gibbon, G.: Reformulation of the theory of conceptual spaces. Information Sciences **177**(21) (2007) 4539–4565
5. Keßler, C.: Conceptual spaces for data descriptions. In: The cognitive approach to modeling environments (CAME), workshop at GIScience. (2006) 29–35
6. Rickard, J.T.: A concept geometry for conceptual spaces. Fuzzy optimization and decision making **5**(4) (2006) 311–329
7. Banaee, H.: From Numerical Sensor Data to Semantic Representations : A Data-driven Approach for Generating Linguistic Descriptions. PhD thesis, Örebro University, School of Science and Technology (2018)
8. Quine, W.V.O.: Ontological relativity and other essays. Number 1. Columbia University Press (1969)
9. Gärdenfors, P.: Induction, conceptual spaces and ai. The Dynamics of Thought (2005) 109–124