Conceptual Spaces: A Bridge Between Neural and Symbolic Representations?

The cognitive framework of conceptual spaces [1] attempts to bridge the gap between symbolic and subsymbolic AI by proposing an intermediate conceptual layer based on geometric representations. A conceptual space is a high-dimensional space spanned by a number of quality dimensions representing interpretable features. Convex regions in this space correspond to concepts. Abstract symbols can be grounded by linking them to concepts in a conceptual space whose dimensions are based on subsymbolic representations.

The framework of conceptual spaces has been highly influential in the last 15 years within cognitive science and cognitive linguistics. It has also sparked considerable research in various subfields of artificial intelligence, ranging from robotics and computer vision over the semantic web and ontology integration to plausible reasoning.

Although this framework provides means for connecting concepts from the symbolic layer to numeric information from the subsymbolic layer, it does not yet provide an automated way to do so: In practical applications, both the mapping from the subsymbolic to the conceptual layer (i.e., how to map observations to points) and the mapping from the conceptual layer to the symbolic layer (i.e., how to map regions to symbols) need to be handcrafted by a human expert.

After introducing the conceptual spaces framework in more detail, I argue that we can use machine learning in order to learn these mappings: I propose to use representation learning (namely the InfoGAN framework [2]) for learning the dimensions of a conceptual space from unlabeled data. Moreover, I suggest to use an incremental clustering algorithm to discover meaningful regions in a conceptual space that can then give rise to abstract symbols.

[1] Peter Gärdenfors, “Conceptual Spaces: The Geometry of Thought”, MIT Press, 2000.

[2] Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel, “InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets”, Advances in Neural Information Processing Systems, 2016.