

D3.3 - AML-DL Specification - Revised

M36

AML learns from data by calculating a subset of the subdirectly irreducible components of an algebraic structure in which the data has been represented.

The process of encoding the data in an algebra is referred to as an algebraic semantic embedding or simply embedding. We use the expression “semantic embedding” as it has been used by mathematicians and theoretical computer scientists for a century and not to refer to representations of text into vector spaces.

Establishing a methodology to build semantic embeddings for real-life problems is challenging. Our detailed manuscript “Semantic Embeddings in Semilattices” explains in detail how to do that for a particular kind of algebra: semilattices. The goal of AML-DL is to provide a set of primitives to write algebraic semantic embeddings. In its current form, AML-DL suffices to write any semilattice embedding.

In this last iteration, we have added the native use of real-valued numeric fields that represent ordered quantities. This extension makes it possible to embed real-valued quantities using a much smaller amount of constants than with the previous version of AML-DL and removing entirely the problem of properly discretizing continuous data.

This specification then moves from being mainly useful for Boolean type values to now becoming fully functional and optimized for numerical values too. This is a very significant extension in terms of usability over the initial AML-DL specification.