Constraint Learning

Attila Kondacs

Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: Constraint learning is machine learning method that finds constraints in the space of temporal event sequences (as in speech), or spatial event sequences (as in vision tasks that involve still images), or spatio-temporal event sequences (as in vision tasks with video streams of images). It learns the internal structure of a space that populates a much larger domain. To make generalizations and fast learning possible, constraint learning makes use of the sparsness of the embeded space as a subspace of the embedding domain. For example all English speech utterances embedded in the space of all possible noise patterns. Once the representation for the constraints has been established, the algorithm, when presented with noisy, partially occluded data, finds the most excited constraint and enforces it. In other words it fills in missing details.

I rephrase the above semi-formally as an example. Assume that we are trying to learn a relation $f \subset A_1 x A_2 x .. x A_n$. We would then present the constraint learner with sequences $(a_1, a_2, a_3, ..., a_n)$ from f to train the algorithm. In the recognition phase similar sequences in which some of the a_i values are missing or incorrect would be fed to the algorithm, The trained constraint learner would fill in the correct a_i.

Motivation: The widely used probabilistic learning algorithms suffer from serious shortcomings when compared to human performance. One of the gravest among these is that recognition performance breaks down ungracefully in the presence of noise and occlusion. Statistical learning algorithms seem to work only under the assumption that the mean of the noise (over space or time) is constant, or is changing so slowly that we can assume it is constant for short periods of time or spans of space, and thus we can measure and estimate it. This is hardly ever true. For example when listening to someone talk, there may be several other louder converstations or noisy passing cars in the background that do not perturb our ability to understand the speaker we are concentrating on; or a significant part of a car can be obstructed from our view by another object and we still recognize the car with ease.

Previous Work: I do not know of any previous efforts in this direction.

Approach: I have been trying to live by the following principles when constructing the constraint learner:

The algorithm should

1. work at several levels of abstraction,
2. build the constraint, representation on its own, i.e. be unsupervised,
3. have uniform structure throughout the hierarchy,
4. be made of simple computational elements or nodes,
5. be able to generalize over the invariances of the domain.

I build on many of the notions explained in [1], and the method itself can be considered as an extreme generalization of that paper.

Research Support: I am supported by an MIT teaching assistantship.

References:

[1] K. Yip G.J. Sussman. Sparse representation for one-shot learning. Technical Report 1633, MIT Artificial Intelligence Laboratory, May 1998.