Appunti del paper *Learning Task-Dependent Distributed Representations by Backpropagation Through Structure*

A.A. 2023/2024

Rosso Carlo

Contents

1 Introduction

The paper introduce a new method to compute and process the backpropagation in complex structured input data. It is particularly interesting for its application with trees and graphs.
We present a connectionist architecture together with a novel supervised learning scheme which is capable of solving inductive inference tasks on complex sumbolic structures of arbitrary size. The most general structures that can be handled are *labeled directed acyclic graphs* (DAGs).
Particularly, we present a simple architecture together with a novel supervised learning scheme that we call *backpropagation through structure* (BPTS). It allows us to generate distributed representations for symbolic structures.

2 Architecture

2.1 Labeling RAAM

The general structure is that of a three-layer feedforward network.
The idea is to obtain a compressed representation for a node of a labeled directed graph by allocating a part of the input of the network to represent the label and the rest to represent its subgraphs. The network is trained by backpropagation in an autoassociative way using the compressed representations recursively. As the representations are consistently updated during the training, the training set is dynamic, starting with randomly chose representations.
Note that this architecture is used to compare the performance with the novel one, the folding architecture.

2.2 Folding architecture

The frist two layers occupy the role of the encoder, the hidden units are connected to a simple sigmoid feedforward layer, in our case just one unit for classification. For the classification of a new structure, the network is virtually unfolded to compute the structure's representation.
We use the vitual unfolding of the network also for the learning phase and propagate the classification error through the whole virtually unfolded network.

3 Back Propagation Through Structure

For both the architecture the number of forward and backward phases per epoch is linear to the number of nodes in the DAG-representation of the training set.

3.1 BPTS for Trees

For reason of simplicity we first restrict our considerations to tree-like structures. In the forward phase the computation of the representation doesn't differ between the two approaches. The following metaphor helps us

to explain the backward phase. Imagine the encoder of the folding architecture is virtually unfolded (with copied weights) according to the tree structure. Now the error passed from the classifier to the hidden layer is propagated throught he unfolded encoder network. Follows the exact formulation.