Development of a strategy for constructing complex solutions by exploiting planning and the usage of machine learning techniques for the resolution of subproblems

MAP-I Doctoral Thesis Proposal for the 2008/09 Edition

1. Context / Background

The task of constructing complex solutions requires that the given problem is decomposed into different subproblems, then a solution is elaborated for each one and finally the different parts are jointed together. We are interested in solutions that exploit machine learning (ML) methods in this process, where the ML provide the solution of individual subproblems.

Consider, for instance, a text extraction system. This system can be composed of various subsystems, some oriented towards, morphosyntactic analysis, tagging, or word sense disambiguation. This may be followed by selection of informative attributes and finally generation of the system for the extraction of the relevant information. Machine Learning techniques may be employed in various stages of this process.

The problem of constructing complex systems can thus be seen as a problem of planning to resolve multiple (possibly interacting) tasks. So, one important issue that needs to be addressed is how these multiple learning processes can be coordinated. Each task is resolved using certain partial order of operations. Meta-learning can be useful in this process. It can help us to retrieve previous solutions conceived in the past and re-use them in new settings.

2. Objectives

The aim of this thesis work is to enhance the existing methodology concerning how to proceed when elaborating a solution to a new complex problem. The method could exploit both retrieving existing solution and adapting them to the current settings (e.g. by re-planning), or by constructing a new solution from scratch, taking into account various type restrictions. Consider, for instance, that output Ox of some algorithm Ax is to be used as input Iy of another algorithm Ay. Then Ox must satisfy the type restrictions required by Ay. The sequence of Ax and Ay is admissible if these type restrictions are satisfied.

The aim of this work is first to explore the possibilities, bring in knowledge gathered in different but related and relevant areas and advance the work in this new research area.

3. Supervision and Proponent Institution

Pavel Brazdil, http://www.liaad.up.pt/~pbrazdil/
LIAAD - INESC Porto L.A., University of Porto
Rua de Ceuta 118 - 6º, 4050-190 Porto
pbrazdil@liaad.up.pt

References

A.Bernstein, F.Provost, and S.Hill, Towards Intelligent Assistance for a Data Mining Process, IEEE Transactions on Knowledge and Data Engineering, 17(4):503-518, 2005.

Brazdil, P. and Bernstein, A., Proceedings of the Planning to Learn Workshop (PlanLearn 2008),

associated with ICML 2008, Helsinki, Finland

Brazdil, P. and Bernstein,A., Proceedings of the Planning to Learn Workshop (PlanLearn 2007), associated with ECML/ PKDD 2007, Warsaw, Poland

Pavel Brazdil, Christophe Giraud-Carrier, Carlos Soares, R. Vilalta, Meta-Learning: Applications to Data Mining, to be published by Springer – 2008

Pavel Brazdil, Carlos Soares, Joaquim P. Costa, Ranking Learning Algorithms: Using IBL and Meta-Learning on Accuracy and Time Results, Machine Learning, Volume 50, Number 3, page 251-277 - March 2003