MALEF
A Framework for Agent-Based Distributed Machine Learning and Data Mining

Jan Tožička1
Michael Rovatsos2
Michal Pěchouček1

1 Agent Technology Group, Gerstner Laboratory,
Czech Technical University in Prague
tozicka,pechoucke@labe.felk.cvut.cz

2 School of Informatics, The University of Edinburgh,
Edinburgh, United Kingdom
rovatsos@inf.ed.ac.uk

Objective: To devise a sufficiently general, abstract view of describing autonomous learning processes in order to be able to utilise the whole range of methods for (i) rational reasoning and (ii) communication and coordination offered by agent technology so as to build effective distributed learning systems.

Generic Model of Learning Process
A single iteration of a learning process can be seen as an agent reasoning cycle that has the following structure:

Can we exploit this to investigate rich forms of interaction between learners in a distributed machine learning system?

Integration Matrix
The generic learner model allows us to look at the different possibilities for information exchange among learners:

<table>
<thead>
<tr>
<th>i</th>
<th>p_i</th>
<th>D_i</th>
<th>H_i</th>
<th>f_i</th>
<th>g_i</th>
<th>h_i</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>p_1</td>
<td>D_1</td>
<td>H_1</td>
<td>f_1</td>
<td>g_1</td>
<td>h_1</td>
</tr>
<tr>
<td>2</td>
<td>p_2</td>
<td>D_2</td>
<td>H_2</td>
<td>f_2</td>
<td>g_2</td>
<td>h_2</td>
</tr>
<tr>
<td>3</td>
<td>p_3</td>
<td>D_3</td>
<td>H_3</td>
<td>f_3</td>
<td>g_3</td>
<td>h_3</td>
</tr>
</tbody>
</table>

Knowledge exchange problem: what knowledge to exchange, when to exchange the knowledge, and how to use received knowledge

Learner Coordination
Evaluation of Example Merging Operators
So far, we have experimented with merging the models/hypotheses of different learners using a contract-net-style approach. Let h_i the receiver’s own model, h_j the provider’s model in a clustering scenario:

\[ p_{h_i \rightarrow h_j}(h_i, h_j) = \]

- **m-join:** The m best clusters (in terms of coverage of D_j) from hypothesis h_i are appended to h_j.
- **m-select:** The set of the m best clusters (in terms of coverage of D_j) from the union h_i ∪ h_j is chosen as a new model. (Unlike m-join this method does not prefer own clusters over others.)

\[ p_{h_i \rightarrow h_j}(h_i, h_j) = \]

- **m-filter:** The m best clusters (as above) from h_i are identified and appended to a new model formed by using those samples not covered by these clusters applying the own learning algorithm f_j.

Experimental Results
Homogeneous Learners
3 k-means agents

Heterogeneous Learners
2 k-means agents and 1 k-medoids agent

Performance: Speed Up
The m-filter operation, decreases the number of learning samples and thus can speed up the learning process.

<table>
<thead>
<tr>
<th>m</th>
<th>filtering</th>
<th>Speedup</th>
</tr>
</thead>
<tbody>
<tr>
<td>30-40 %</td>
<td>10-20 %</td>
<td></td>
</tr>
<tr>
<td>20-30 %</td>
<td>5-15 %</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion
Even a very simplistic application of MALEF architecture has proven capable of improving the performance of individual learning agents.

In the future we want to explore other types of combining the different elements of different learners through elaborate communication and reasoning methods.

Acknowledgement
We gratefully acknowledge the support of the presented research by Army Research Laboratory project N62558-03-0819 and Office for Naval Research project N00014-06-1-0232.