Chapter 8

Conclusions

The work on our research followed three different directions. On one side, it was focused on developing a framework with enough flexibility to allow us to modify components and automatically generate an Agent System. On another side, it was focused on analyzing Agent Systems on decision problems and providing more insight on useful techniques over these systems. Finally, it was also focused on using the knowledge gathered of Agent Systems to solve a novel problem which had not been solved before through learning methods. These three paths converged in the end of the research, producing a framework, which is able to face new learning challenges, with which multi-agent system components where analyzed and a complete multi-classifier for the HLA classification problem was generated.

Mage was the name assigned to the framework, which stands for Multi-Agent Generation System. Its architecture was based on a three layer design: The framework, the implementation layer and the application layer. The design allowed the implementation of different components for the preprocess, partition, train and postprocess stages of an Agent System generation process. Through the graphical application, MageGui, it was possible for the user to choose the components he wanted in the generation process, and have the framework execute the different stages, up to the point where the resulting

agent system was exported into an XML representation, which could be deserialized by any other application, regardless of the language or implementation.

Using the Mage framework, a set of experiments was performed, with the specific purpose of evaluating different preprocess and learning components in the process. Soft and Hard partition were compared, and the obtained results showed the close relationship between the sample features and the partition method which produced best results. Adaboost, Marginal Adaboost and Adaboost* were comparatively analyzed, and it was shown empirically how Adaboost* was a superior technique against these other boosting methods. The properties of bagging were verified through experimentation as well, allowing us to provide empirical evidence to the statistical problem solution that Agent Systems provide. Finally, the implementation correctness was as well verified through experiments which compared results against benchmarks.

Special focus was placed on the HLA problem. First, the problem was translated from an optimization to a classification problem, mapping it to a multi-classification task similar to classifying handwritten digits. Once defined the problem, the components were selected according to the knowledge gathered from previous experiments: Hard Partitioning and Adaboost* were used as Partition and Learning components respectively. The generation of the multi-classifier took around 2 months, but in the end a multi-classifier was generated that topped the expectations by achieving 97 % correct predictions. This classifier was exported and used in a simple external application which now has the power of classifying HLA, and hence achieving the goal of, through the framework, allowing any application to include multi-agent systems easily. The results of the multi-classifier gave us also insight on the advantages and disadvantages of the technique, and its applications further from classification, as are knowledge representation and dataset analysis.

In the end, the three goals were reached, as now we have a system that will be a basis

for further research and development on agent systems for decision making problems, a reference for Multi-Agent system characteristics, and a system which will allow further analysis of the HLA classification problem as it continues to be studied using either machine learning or other techniques.

8.1. Future Work

The work in our research gave us insight on which components in each stage are useful, based on the characteristics of the problem at hand. From our experimentation results, we obtained evidence that in the case of binary classification, when there is a high variation in the features of different samples, hard partition is more useful, but when samples are similar, soft partition performs better. We also learned that Adaboost* is superior than Marginal Adaboost and Adaboost, and that we may calculate the number of iterations required for maximizing the margin. We learned that in the preprocess stage, there are common methodologies for removing unnecessary features. We learned that increasing the number of agents does not necessarily imply better performance. Finally, we learned that a classification problem may be considered prediction, and that we may identify one case from the other based on the nature of the output.

All the information we discovered in experimentation, which was applied when designing the HLA multi-classifier, is useful for proposing the user a set of components to be used for the creation of the new agent system. Through measures like variation in sample features, dataset size, number of outputs and nature of features (binary values, integers, real numbers), we may create a description of the dataset, which together with the Metaknowledge gathered in our research, will allow us to completely automate the agent system creation process.