Explorative Analysis
and Visualization of
Large Information Spaces

Learning Hierarchical Rule Systems

Thomas Gabriel

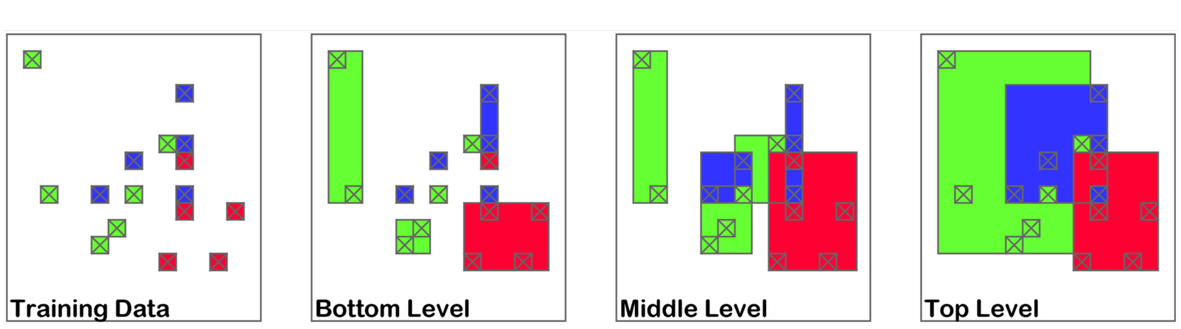
Universität Konstanz

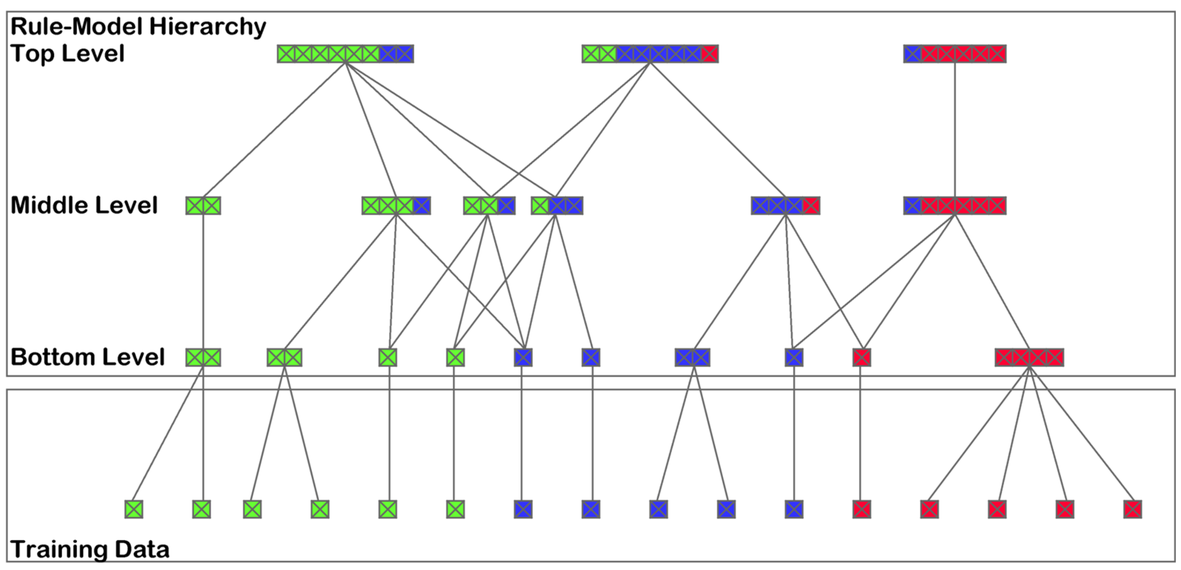
1 Motivation

Learning rule systems from real-world data often generate too many rules for human interpretation. Hierarchical rule systems with only a small number of rules at each level can help to overcome this issue: Lower levels in this hierarchy focus on areas of the feature space where only weak evidence for a rule is found in the data. Rules further up, at higher levels of the hierarchy, describe increasingly general and strongly supported aspects of the data. Based on a hierarchical extension of a base fuzzy rule learner, we show a better classification accuracy where at the same time the number of rules decreases. This approach leads to a well-defined arrangement of hierarchy level and allowing visual interactions across the different rule models. We demonstrate how a connected set of parallel coordinates can be used to visually explore this hierarchy of rule systems and enables an intuitive mechanism to zoom in and out of the underlying model hierarchy. A second approach for visualizing hierarchical rule models known as multi-dimensional scaling, is applied to map rule centers, corresponding coverages, and data onto one coherent 2D plot.

2 Hierarchical Rule Systems

The figure below shows a 3-level rule hierarchy where the bottom level contains non-conflicting rules (most specific) covering only instances of the same class, the middle level (few conflicting patterns), and the top level (most general) has one rule per class.

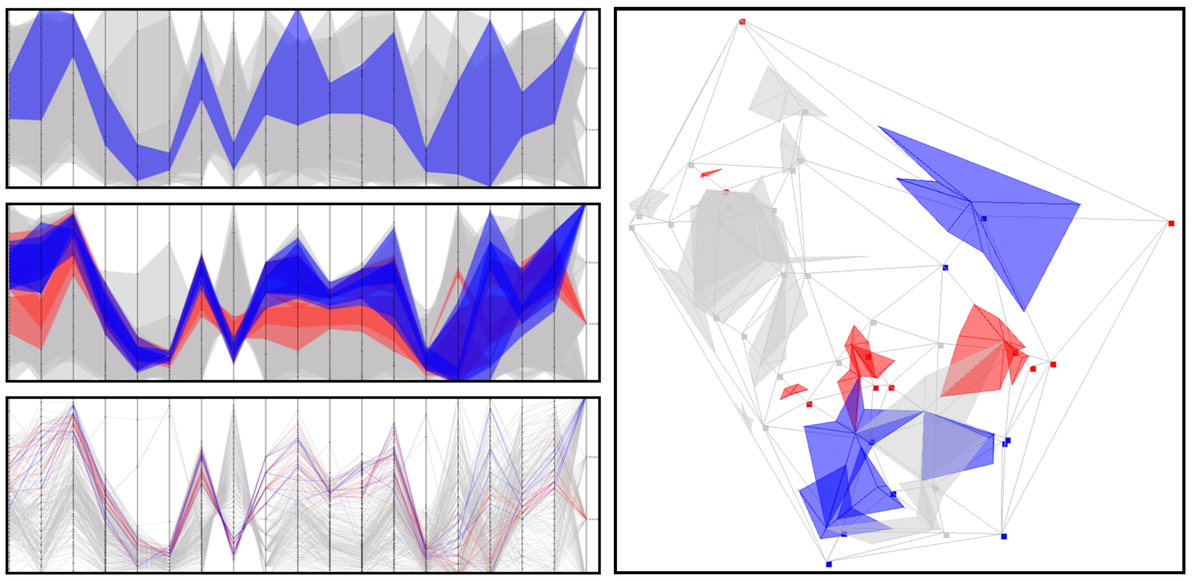




The underlying rule algorithm along with an intelligent filter strategy allows the generation of such a hierarchy. The rule layers are arranged in a hierarchy of different levels of precision. The rule induction algorithm is used to determine rules of low relevance, which are then used as a filter for the next training phase. Training examples that resulted in the creation of small, less important rules are therefore excluded from the training phase of the next layer, resulting in a more general rule system, ignoring the small withheld details in the training data.

3 Visualizations

Parallel Coordinates are used to display rules as connected intervals, shown as colored bands. Each level is visualized independently and allows interactions across different levels by depicting rules in one view. A 3-level rule hierarchy is generated on the Vehicle Silhouettes data. Here, one selected rule in the top level results in a number of selected and overlapping rules in the levels below, which show more specific rules of different classes and rules explaining outliers and artifacts in the bottom level.
Multi-Dimensional Scaling is used to display the hierarchical rule model on the Vehicle Silhouettes data. The rule centers and data points are mapped onto one coherent picture.



4 Results and Conclusions

The generated hierarchy of rule systems provides more accurate results compare to non-hierarchical algorithms, and at the same time uses less rules which allow visual exploration across the different levels of the hierarchy.

References

[1] Th. R. Gabriel and M. R. Berthold, *Influence of fuzzy norms and other heuristics on "Mixed Fuzzy Rule Formation"*, International Journal of Approximate Reasoning (IJAR), Elsevier, 2004, 35:195–202.

[2] Th. R. Gabriel and M. R. Berthold, *Missing Values in Fuzzy Rule Induction*, IEEE Conference on Systems, Man and Cybernetics (IEEE SMC), IEEE Press, 2005.

[3] Th. R. Gabriel, A. S. Pintilie, and M. R. Berthold, *Exploring Hierarchical Rule Systems in Parallel Coordinates*, Proc. 6th International Symposium on Intelligent Data Analysis (IDA), Lecture Notes in Computer Science 3646 (LNCS), Springer Verlag, pp. 97–108, Madrid, 2005.

[4] Th. R. Gabriel, K. Thiel, and M. R. Berthold, *Rule Visualization based on Multi-Dimensional Scaling*, IEEE International Conference on Fuzzy Systems (Vancouver, BC, Canada), 2006.