Learning Relational Structures for 2D Pattern and 3D Object Recognition

Adrian R. Pearce[1], Terry Caelli[1] and Walter F. Bischof[2]
Email: adrianp@cs.mu.oz.au, tmc@cs.mu.oz.au
wfb@psych.ualberta.ca
Fax: +613 282 2490

[1]Department of Computer Science
The University of Melbourne
Parkville, Victoria, 3052, Australia

[2]Psychology BSP-577
University of Alberta
Edmonton, Alberta,T6G 2E9, Canada

May 15, 2007

Extended Abstract

Over the past fifty years, many different techniques have been developed for recognizing 2D patterns and 3D objects invariant to different types of geometric transformations, contextual conditions and task demands. The representation used will determine the kind of queries which can be made. In Computer Vision, for example, relational structures are particularly suitable for representing complex patterns and providing matching solutions. To this date, however, little attention has been devoted to the problem of developing techniques for the automated learning of such relational descriptions - and

to the associated problem of generalization. This paper is concerned with these issues.

Based upon a fundamental pattern representation in terms of parts and relations, our approach integrates classical techniques from Statistical Reasoning (Evidence Theory) with those from Symbolic Reasoning (Machine Learning). The key to combining these two modes of reasoning is label-compatibility, i.e. the representation must maintain compatibility information (labels) between rule-based description of patterns where rules can describe parts (unary rules) or part relations (binary rules).

In this paper we compare three new techniques for learning relational structure in 2D patterns and 3D objects:

- Evidence-Based Systems (EBS) [2],
- Conditional Rule Generation (CRG) [1],
- and Rulegraphs [3].

Two major problems have emerged in all traditional solutions to the Graph Matching problem. First, the computational complexity is exponential and this is a significant problem since the cardinality of such algorithms is defined by the number of models and sample parts. Second, Feature Indexing has traditionally been model-based in that prior knowledge about class data has been provided. That is, it has not been clear how to learn model descriptions. This is a particular problem when patterns are partially occluded or distorted, as is often the case in pattern or object recognition. Specifically, past algorithms have not considered generalization, i.e. the ability for the system to recognize equivalences between objects which are not identical.

In Statistical Reasoning, we can use techniques from Evidence Theory which are typically used in Expert Systems. Models are represented via the enumeration of features which evidence pattern classes and evidence weights are assigned according to the occurrence of feature states in observed data.

Symbolic Reasoning manipulates relational information using first order inductive logic: Horn clauses. Such systems utilize both conjunctive and disjunctive clauses and can generalize over relational training instances (for example, FOIL [4]).

In contrast, Evidence-Based Systems (EBS) have the ability to generalize pattern rules of the *if-then-else* form over feature space regions (volumes).

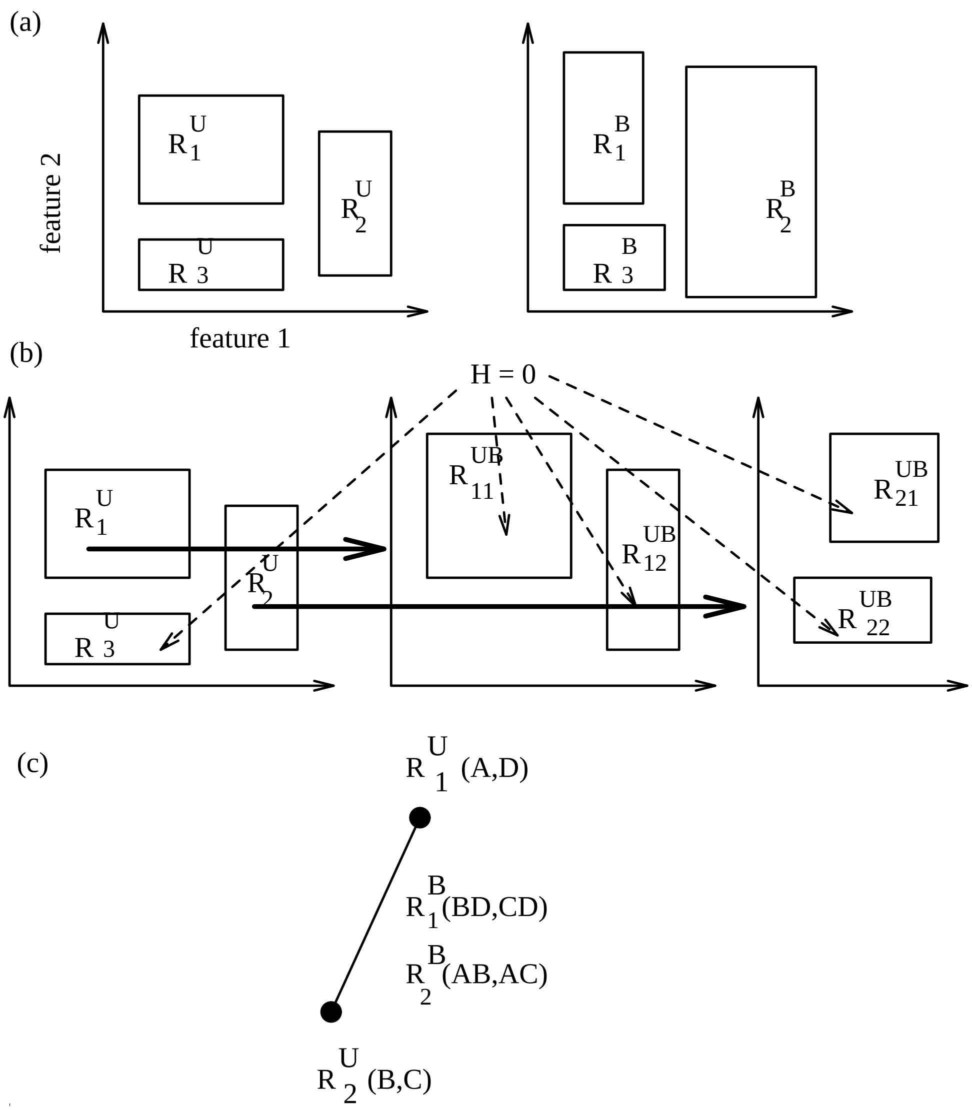

Figure 1: (a) EBS rules are shown for both unary and binary feature spaces and provide both conjunction and disjunction. (b) The Conditional Rule Generation Procedure (CRG) relies on label-compatible paths ($U_i - B_{ij} - U_j - \ldots$). The unresolved clusters are expanded until either all rules are resolved or the predetermined maximum rule length is reached, in which case rule splitting occurs (H refers to cluster entropy). (c) A Rulegraph is shown in which the label-compatibilities between rules is maintained. The cardinality is reduced with respect to the number of parts in the training instances.

Conjunction and disjunction is allowed using bounding boxes (rules) and minimum entropy or distance cost functions are used to optimize class discrimination. Rules are generated over both unary (single parts) and binary (part relation) numerical features. (see Figure 1a). Label-compatibility is implicitly encoded by calculating evidence weights for non-linear combinations of unary and binary rules using a neural network.

Conditional Rule Generation (CRG) investigates the use of first order inductive logic via union and discrimination trees (see Figure 1b). In particular, CRG takes into account the explicit label-compatibilities which should occur between unary and binary rules - in their very generation. The CRG technique creates cluster trees of hierarchically organized rules for classifying structural pattern descriptions which aims at best generalizations of the rule bounds. The algorithm searches for the occurrence of more unary and/or binary feature states between connected components of the training patterns. Here, the occurrence of more than one class sample, as measured by a non-zero entropy (information H) statistic, indicates that higher order rules are required. A backtracking rule-splitting procedure resolves ambiguity while achieving greatest generalization.

Rulegraphs apply Machine Induction to the Traditional Graph Matching problem. Here, the label-compatibility of the rules is checked after rule generation. The technique relies on two simple principles: First, sets of model graphs and their vertices are reduced by generalization while maintaining label-compatibility between rules. Second, search is constrained using evidence weights produced by an EBS system from training data. Since the matching process involves graphs of cardinality no greater than the number of unary rules, the process is more efficient than classical Graph Matching procedures (see rulegraph in Figure 1c). In addition, reducing the search space using a breadth first search (such as A* search) has previously not been possible when large numbers of parts or large branching factors are involved.

The systems were tested and compared using a number of different 2D pattern and 3D object recognition problems, with particular emphasis on their suitability for different types of data. Finally, we show how the techniques apply well to the recognition of many different objects in a scene and the extent to which the learned rules can identify patterns and objects which have undergone non-rigid distortions. The techniques are shown to improve the uniqueness of the matching process while reducing the computational complexity.

References

[1] Bischof, W. F., and Caelli, T. Learning structural descriptions of patterns: A new technique for conditional clustering and rule generation. *Submitted for publication* (1993).

[2] Caelli, T., and Drier, A. Variations on the evidence-based object recognition theme. In *Proceedings of the eleventh IAPRA International Conference on Pattern Recognition* (1992), pp. 450 – 454.

[3] Pearce, A., Caelli, T., and Bischof, W. F. Rulegraphs for graph matching in pattern recognition. *Submitted for publication* (1993).

[4] Quinlan, J. R. Learning logical definitions from relations. *Machine Learning 5* (1990), 239–266.