Statement of Interest

John Yen

May 5, 1989

I am interested in attending the workshop and discussing the role of term
subsumption language in rule-based systems. The CONSUL system was the
first to use a term subsumption language (i.e., NIKL) and its classifier in
a rule-based system. It used KL-ONE classification as an indexing scheme
for rules. Similar ideas have been used in other systems such as CAKE and
KL-TWO. However, the inference architecture and rule language of these
systems are not as general as that of a production system. We have thus
designed and implemented a production system using LOOM's term sub-
sumption language. Two major contributions of our work are (1) combining
KL-ONE realization with RETE pattern matching algorithm for performing
multi-variable semantic-based pattern matching, and (2) classifying multi-
variable patterns into a subsumption lattice for computing the speci-
ficity of rules. Although the production system was imple-
mented in LOOM, our results can be easily applied to other term subsumption systems.

My current interests include the following.

1. Explore the trade off between completeness of the pat-
tern classifier and the expressiveness of the rule language and the term subsumption
language. Answers to the following questions are of interest. Whether
the pattern classifier is complete for a complete term sub-
sumption language (e.g., CLASSIC)? Whether the pattern classifier for LOOM
is complete if LOOM has a complete classifier? How disjunction and
negation in rule’s conditions affect the completeness of the pattern classif-
ier?

2. Demonstrate the benefits of pattern classification using rule-based ap-
lications. We have identified at least three potential benefits of the
pattern classifier: (1) It facilitates control of rules because it allows conflict resolution to be made based on a principled specificity relations among rules. (2) It enhances maintainability of rule-based systems because the predictability of the system is increased by an explicitly representation of the rules specificities.

3. It enables a meta-level control language to refer to a class of rules using their common subpatterns. We plan to demonstrate these benefits using empirical evidence.

4. Performance evaluation. I am also interested in the performance of our production system compared to OPS5 and other existing production systems. Two measures that are important for evaluating our system are the computation time for compiling a rule base, and the computation time spent during pattern matching. These two measures are particularly important to see whether our system is ready for large scale applications.

My potential contributions to the workshop are the following.

1. Present the design and the implementation of the pattern classifier. Describe the two major steps of the pattern classification algorithm: (1) normalizing patterns into a canonical form, and (2) searching for a subsumption substitution. The subsumption of patterns is defined as follows: a pattern P1 subsumes P2 if and only if there exists a substitution S such that

\[ P2 \Rightarrow P1/S \]

where S (called the subsumption substitution) is a substitution that replaces each variable in P1 by a variable or a constant in P2. The normalization process deduces implicit conditions (clauses) implied by the definitional knowledge of the pattern’s predicates. The subsumption relation between two normalized patterns is then computed by searching for a subsumption substitution using dependency-directed backtracking. The search space can often be pruned significantly using knowledge about the subsumption relation between predicates. In addition to describing the pattern classification algorithm, I also hope to present our results on the trade off between completeness of the
pattern classifier and the expressiveness of the term subsumption language used. Discuss how different term subsumption language affects the pattern classification algorithm. Discuss the relationship between classifying patterns and classifying N-ary relations.

2. Present the design and implementation of CONCRETE (CONcept Classification RETE). Discuss issues regarding the integration of realization and RETE match algorithm. In particular, how did we define the interface between the two such that caches in the RETE network are consistent with ABox states? How do we avoid duplicating ABox information in the RETE net? How do we reduce the size of RETE caches using efficient ABox retrieval capabilities?

3. Inspire exchange of ideas on the use of term subsumption language in alleviating problems of rule-based paradigm (e.g., control, maintainability, understandability). This effort, I believe, will shed light on many important issues of pure rule-based system and will encourage more AI researchers to consider using KL-ONE family languages as the underlying knowledge representation of their rule systems.

Reference

John Yen, Robert Neches, and Robert MacGregor "Using terminological models to enhance the rule-based paradigm".