Transformation Rules for Knowledge-Based Pattern Matching*

Peter Yeh, Bruce Porter, and Ken Barker
Department of Computer Sciences
University of Texas at Austin
Austin, TX 78712 USA
{pzyeh,porter,kbarker}@cs.utexas.edu

Abstract

Many AI tasks require determining whether two knowledge representations encode the same knowledge. For example, rule-based classification requires matching rule antecedents with working memory; information retrieval requires matching queries with documents; and some knowledge-acquisition tasks require matching new information with already encoded knowledge to expand upon and debug both of them. Solving this *matching problem* is hard because representations may encode the same content but differ substantially in form.

Previous approaches to this problem have used either syntactic measures, such as graph edit distance, or semantic knowledge to determine the "distance" between two representations. Although semantic approaches outperform syntactic ones, previous research has focused primarily on the use of taxonomic knowledge. As a result, mismatches between representations go largely unaddressed.

In this paper, we investigate whether semantic approaches can be augmented with additional non-taxonomic knowledge to further improve matching. To test this hypothesis, we built a matcher that uses both taxonomic and non-taxonomic knowledge in the form of transformation rules and applied it to the task of critiquing military Courses of Action. We compared our matcher's performance to both a syntactic and a semantic matcher applied to the same task. From this study, we found the results show that using additional non-taxonomic knowledge further improves matching.

1 Introduction

A requirement common to many AI tasks is determining whether two knowledge representations, encoded using the same ontology, encode the same knowledge. For example, rule-based classification requires matching rule antecedents with working memory; information retrieval requires matching queries with documents; and some knowledge acquisition tasks require matching new information with already encoded knowledge to expand upon and debug both of them.

The task of determining whether two representations encode the same knowledge is often threated as a graph matching problem. The representations are encoded as either attributed relational graphs or conceptual graphs. Two representations encode the same knowledge (i.e. match) if their corresponding graphs match. The criteria used to decide if two graphs match can be any one of the following: graph isomorphism, subgraph isomorphism, maximal common subgraph, etc.. We call this the *matching problem*.

Solving this *matching problem* is hard because multiple encodings of the same knowledge rarely match exactly, so a matcher must be flexible to avoid a high rate of false-negatives. However, a matcher that is too flexible can suffer from a high rate of false-positives. The problem has various causes, including:

*A shorter version or this paper is available as "Using Transformations to Improve Semantic Matching" which appeared in the proceedings of K-CAP'03.