GOAL-DRIVEN SIMILARITY ASSESSMENT*

Dietmar Janetzko\(^1\) and Stefan Wess\(^2\) and Erica Melis\(^3\)

\(^1\) Dept. of Cognitive Science University of Freiburg
\(^2\) Dept. of Computer Science University of Kaiserslautern
\(^3\) Dept. of Computer Science University of Saarbrücken

Abstract. While most approaches to similarity assessment are oblivious of knowledge and goals, there is ample evidence that these elements of problem solving play an important role in similarity judgements. This paper is concerned with an approach for integrating assessment of similarity into a framework of problem solving that embodies central notions of problem solving like goals, knowledge and learning. We review empirical findings that unravel characteristics of similarity assessment most of which have not been covered by purely syntactic models of similarity. A formal account of similarity assessment that allows for the integration of central ideas of problem solving is developed. Given a goal and a domain theory, an appropriate perspective is taken that brings into focus only goal-relevant features of a problem description as input to similarity assessment.

1 Introduction

In recent years, there has been an upsurge of interest in case-based reasoning (CBR), i.e. reasoning techniques that are based on the use and reuse of previous problem solving experience [Kol91]. One of the key issues of case-based reasoning (Fig. 1) is the question how a previous case, i.e. a source, is selected given a current case, i.e. a target. This retrieval step calls for estimating similarity between source and target cases. The majority of previous approaches to similarity assessment resort to measures of similarity that have been developed within the province of categorization and clustering (e.g. in biology [Dic45]), but not within the realm of problem solving. These approaches have been termed syntactic, as they confine similarity assessment to the objects given in the problem description and refrain from using purposes or goals. In contrast, these factors on the side of the problem-solver are at the heart of the so-called pragmatic approaches (e.g. [Hol85]) to similarity. We take the view of similarity as a genuine part of problem solving. Within this conceptual framework similarity assessment is influenced by two types of characteristics: syntactic characteristics, e.g. number of common features, on the side of the objects of similarity assessment and

* This research was supported by the "Deutsche Forschungsgemeinschaft" (DFG), "Sonderforschungsbereich" (SFB) 314: "Artificial Intelligence and Knowledge-Based Systems", projects X9 and D3.
pragmatic characteristics, e.g. goals, on the side of the subject of similarity assessment.

Fig. 1. Basic steps of CBR

The aim of this article is to develop a model that links pragmatic and syntactic approaches to similarity. The model we propose does not give priority to any of the two accounts on similarity assessment. It is, however, motivated by the fact that the quality of similarity assessment heavily depends on focusing on relevant features. Applying a syntactic approach implicitly requires that relevant features are known before assessing similarity. In contrast, this implicit assumption becomes an explicit part of goal-driven similarity assessment, such that the relevance of features is determined by a knowledge-based process. Reduced to its kernel, our model (Fig. 2) starts with a pragmatic account using a problem solving goal and a domain theory. Following this, a perspective (cf. [Str91]) is developed which brings into focus features that are important for goal-achievement. Finally, similarity of a given object to other objects can be computed by applying a syntactic similarity measure e.g. [Tve77] to the goal-relevant features.

Fig. 2. Goal-driven similarity assessment