Semantic Virtual Engineering

**PROJECT OBJECTIVES**

The SEVENPRO project aims at developing technologies for:
- mining of product engineering knowledge from multimedia repositories and
- semantically enhanced 3D interaction with that knowledge in integrated engineering environments through virtual reality.

CAD designs, documental repositories and enterprise-resource-planning databases will be the main data/knowledge sources supported.

The project brings together the so far distinct worlds of 3D design data and other representation formats of engineering knowledge.

**DESIGNS ANNOTATION**

The information available in CAD files and other data sources is formalized and integrated by means of semantic annotation based on ontologies. Semantic annotation of CAD designs
- is generated automatically from the commands history available via the API of CAD tools
- is based on a CAD ontology developed in SEVENPRO and available in RDF format

Annotation including ontology of CAD items and axioms defining core relations is automatically translated into Prolog.

Relational Mining from Product Designs Data

**MINING ENGINEERING DESIGNS**

Discovery of interesting design patterns from CAD repositories is an expected aid to reusing design knowledge. Engineering designs contain implicit expert knowledge. The data mining algorithms will help to reveal and make explicit. Such obtained knowledge can be reused as corporate design standards or recommendations: to support the work of engineers (reusability), to check pattern compliance of new designs (quality checking), to teach novel engineers on how to design specific parts (training).

**OUR APPROACH**

Our approach is based on the sorted refinement operator, i.e. a subsumption relation combining classical sub-subsumption with taxonomies on terms and predicates.

Propositionalization enables us to exploit the whole variety of available data mining algorithms. Propositionalized representation of classified relational data is generated by constructing first-order features. During the feature generation a table of mutual feature subsumptions is maintained. This subsumption is exploited in propositional pattern search, which prunes any conjunctions of a subsumer with its subsumer and specializes a conjunction not only by extending it, but also by replacing an included feature with its subsumer.

Finding maximal patterns For non-classified data, find maximal patterns of some maximal length covering the minimum set amount of examples.

**RESULTS AND ONGOING WORK**

In a proof-of-concept experiment, the system was tested on a database of 33 CAD designs. One design usually contains around 100 properties. It is not feasible to generate features with tens of literals, therefore the maximum length of a feature was set, based on maximum depth and highest occurrence of one property in the dataset. An exemplary maximal pattern is below:

Currently we are struggling to improve efficiency by use of graph search techniques for feature generation. For principle exploitation of hierarchical background knowledge, we are also considering approaches using hybrid languages such as AL-log or using a more complex description language such as $\gamma$-terms or antecedent description grammars.

This research is supported by SEVENPRO: Semantic Virtual Engineering for Product Design, project IST-027473 (2006-2008) funded under 6th Framework Programme of the European Commission.

The 16th International Conference on Inductive Logic Programming (ILP 2006), Santiago de Compostela, Spain, August 24-27, 2006