Knowledge Acquisition and Visualization for Biomedical Research

Noah Zimmerman, BS\(^1\), Stephen Meehan, BS\(^2\), Leonore A. Herzenberg, PhD\(^3\),
Amar K. Das, MD, PhD\(^1\)

\(^1\)Stanford University, Biomedical Informatics; \(^2\)ScienceXperts Inc, Palo Alto, CA; \(^3\)Stanford University, Department of Genetics

Abstract

The TreeTableWidget is a user interface component that allows instances in a knowledge-base to be sorted on the value of one or more of its attributes. Any combination of attributes can be sorted simultaneously by assigning a precedence to each sort. This results in a hierarchy of attributes such that instances can be grouped and visualized according to shared attribute values in a standard tree view. The resulting tree provides a mechanism to navigate and edit a large number of instances based on attribute values. The component is available as a Protégé slot-widget plug-in.

Introduction

Symbolic representations of the knowledge relevant to biomedical research is an area of active development\(^1\). Ontology is one type of representation framework that has gained widespread acceptance in the biomedical community. An ontology, in this setting, is a formal way of representing the concepts and relationships relevant to a domain.

One of the challenges associated with developing and interacting with complex data structures is providing a user interface that conveys the underlying model clearly and concisely in different contexts. We have developed a reusable user interface component for the Protégé platform for visualizing and interacting with a knowledge-base. The component is designed for displaying large numbers of classes and instances with many properties by providing customizable views by which the user can organize the knowledge-base.

Background

Protégé\(^3\) is an open source ontology editor and knowledge-base framework that can be used to develop ontologies and perform knowledge acquisition.

We have used Protégé as a platform for developing CytoGenie™\(^4\), a knowledge-based application to support researchers as they plan and design a flow cytometry experiment. One of the core components of flow cytometry experiments are reagents, which are used to label cells of interest. Clinical and research laboratories frequently have several hundred reagents in inventory. The software guides the user in selecting a feasible combination of reagents for an assay based on experimental conditions.

Description

The primary requirement for the widget is facilitating knowledge acquisition and presentation by domain experts with no background in formal knowledge representation. A user must be able to navigate thousands of instances by groups of related concepts, modify members of a group, display information differently for each context, and integrate with Protégé.

The TreeTableWidget (TTW) is a Protégé user interface that allows knowledge-base instances to be sorted on the value of one or more of its attributes. Furthermore, the attribute sort order can be ranked. The ranked sort order creates a hierarchy of attributes such that instances can be grouped and visualized in a standard tree view. The resulting tree provides a mechanism to navigate and edit a large number of instances based on shared attribute values.

The component has been used to represent the reagent inventory for individual labs as well as reagent manufacturers.

\(^1\) Payne PR, Deitzer JR, Mendonca EA, Starren JB. "Consensus-based Construction of a Taxonomy of Clinical Trial Tasks." AMIA Annu Symp Proc. 2006;
\(^4\) www.ScienceXperts.com