Visualizing Semantic Wiki Structure **Exploring the Hesperian Digital Library**

Problem

Semantic wikis provide rich formal structure for content, but current software does not present this information in an accessible manner. How can such information be presented to reveal relationships and facilitate navigation?

Motivation

A semantic wiki is a wiki which utilizes an ontology to represent formalized knowledge. Current semantic wiki software leverages this representation primarily to support enhanced queries.



Semantic wikis offer a simple syntax for annotating formal properties (top). These are extracted (botand made available tom) for queries.

A space-filling radial layout is used to represent the hierarchical structure of diseases. The ordering is done such There are existing ontology visualization tools, but they that diseases with similar symptoms are close together. are ill-suited for this application since they are discon-Element sizes are scaled in proportion to the amount of nected from the article content of the wiki. Presenting a information about them in the ontology. visualization alongside the wiki in the browser allows it be used as a navigation tool for someone exploring the site. Users can interactively manipulate the visualization. To focus in on a portion of the graph, they can remove a sub-By constructing a visualization tailored to the particular tree from its parent to create a separate sunburst. characteristics of the domain, it is possible to derive useful

information that is only implicit in the data set. In the case There is bidirectional communication between the visualiof the Hesperian Digital Library's healthcare domain, it zation and the wiki. Clicking on nodes in the visualization would be useful to determine which diseases have similar loads the corresponding page in the wiki and navigating causes, symptoms, or treatments. Such similarity calculathough the wiki highlights the corresponding node in the tions can be used to constrain the layout of a visualization. visualization.

Approach

Users can click nodes in ren's Disease: Earache and Ear Infection the visualization to navigate to the corresponding page hesperian Earache and Ear Infections in wiki. Publishing for community ealth and empowerme Ear infections are common in small children. The Main Page nfection often begins after a few days with a cold or a Recent chang tuffy or plugged nose. The fever may rise, and the child Random page ften cries or rubs the side of his head. Sometimes pus Ontology Bro the ear. In small children an ear infection ometimes causes vomiting or diarrhea. So when a child as diarrhea and fever be sure to check his ears. All articles Disease sur It is important to treat ear infections early. Give an antibiotic like penicillin (p. 351) or cotrimoxazole (p. 358). In children under 3 years of age, ampicillin (p. 353) often works better. Give acetaminophen (p. 380) (Go) (for pain. Aspirin also works but is less safe (see p. 379). Carefully clean pus out of the ear with cotton, but do not put a plug of cotton, a What links here stick, leaves, or anything else in the ear. Related chan Children with pus coming from an ear should bathe regularly but should not swim or dive for at least 2 weeks after they are well. 1. Grab Category: Common Sickness

Symptom Similarity For each pair of diseases described on the Hesperian wiki we compare their sets of symptoms. The symptom property has string values, such as "Severe coughing" and "Mild fever" so simple natural language pre-processing is applied. After stemming and removing stop-words, we compute a score based on the number of tokens in common (Jaccard index). From these pair-wise property scores, we select the assignment that maximizes the overall similarity of the two diseases. Layout Ordering We seek a radial layout in which similar diseases are arranged close together. This objective corresponds to minimizing a cost function which incorporates the similarity scores and the radial distance between nodes. Due to the computational complexity of finding the optimal ar-2. Drag rangement, simulated annealing is used to obtain an ap-Users can drag subtrees to split them off into independent sunbursts. proximation.

Results

CS 294–10: Visualization UC Berkeley & Fall 2008 Matt Gedigian

Future Work

Category: Common Sickness

Common Sickness: Cough

3. Release

This work demonstrates the potential for integrating interactive visualizations with semantic wikis. This visualization could be improved by:

- including node labels
- incorporating search and filter controls
- allowing the user to select from different similarity metrics



