SemGrail 2007

June 21 – 22, 2007

Marriott Redmond Town Center

Redmond, WA

Contents

Consensus Data, Ontologies, and Services

Michael Huhns (University of South Carolina)

There is a power in consensus that is not being exploited sufficiently by Web applications and Web users. The goal of my research is to identify the various problems for which consensus is an appropriate solution strategy and then to develop the strategies. The challenges will be to model the semantics of the Web pages involved, to reconcile the semantics of the large number of pages needed to form a consensus, to model the context of a domain problem or user interest, and to reach a consensus that solves the domain problem in the appropriate context. The results will be embodied in publishing, search, and domain-specific problem-solving tools. Ultimately, in an open environment of numerous, heterogeneous services, information sources, agents, and resources (processors, storage, bandwidth), users and developers would be able to achieve their desired outcomes by specifying what, not how the outcomes should be achieved.

Example 1. "What is the population of China?" My personal agent could solve this, with the help of Google, if (1) each Web page is marked-up with semantic tags; (2) there is a common ontology of tags or my agent is able to reconcile disparate ontologies; (3) my agent understands the context of my question, that is: I want the current population, not the one in 1800, and I want the resident population, not residents + tourists, etc.; and (4) when a Google query returns 70M pages, my agent could combine the data from many of them, possibly by averaging or by choosing the most common value (i.e., a vote).

Example 2. "If current trends continue, when will the population of India equal the population of China?" All of the information needed to answer this automatically is available, but this is a much harder problem than Example 1, because it requires forming a consensus on trends and then extrapolating.

Example 3. "Which is the correct grammar: "in the West Coast," "on the West Coast," or "at the West Coast?" In Google, the first yields 0.6M hits, the second 10.2M hits, and the third 0.2M hits. In essence, Google can be used to disambiguate grammar by forming a consensus.

Example 4. "How can two sources each having semantics specified independently be reconciled or enabled to interact without confusion?" If the semantics of numerous other sources on the Web can be used to form a consensus over the semantics of the two sources, reconciliation might be achieved.

Example 5. "What is the best Web service for providing stock quotes?" Given that there are many Web services available for providing stock quotes, the best one can be chosen by finding which provides the consensus result most often. Alternatively, the functionality of all Web services can be combined to produce a service better than any of the individual services.

Example 6. "What is the best, or consensus, recipe for baking chocolate chip cookies?" A Google search returns 1.1M hits, but theory is needed for how recipes, i.e., plans or workflows, can be combined.

Example 7. "Who programs the lights at a highway intersection?" That is, all of the cars waiting at a light could decide dynamically on the operation of the light by forming a consensus. But, how can consistent and fair operation be achieved?

A research program is needed to explore the requirements for and feasibility of a range of strategies for reaching and applying the results of consensus. My team is investigating a variety of tools, based on multiagent systems, that can form a consensus about data, semantics, functionality, and behavior.

Dr. Michael N. Huhns is the NCR Professor of Computer Science and Engineering and director of the Center for Information Technology at the University of South Carolina. His degrees in electrical engineering are from the University of Michigan (B.S.) and the University of Southern California (M.S. and Ph.D.). He is the author of seven books and more than 200 papers in machine intelligence, including the recently coauthored textbook Service-Oriented Computing: Semantics, Processes, Agents. He serves on the editorial boards for 8 journals, is a Senior Member of the ACM, and is a Fellow of the IEEE.