Argos: Dynamic Composition of Web Services for Goods Movement Analysis and Planning

Jose Luis Ambite[1], Genevieve Giuliano[2], Peter Gordon,
Andreas Harth, LanLan Wang, Qisheng Pan , and Stefan Decker

This research will develop a flexible data query and analysis system based on the web services paradigm. As an application domain we will examine several goods movement planning problems and their effects on urban structure. The project began in August 2003, and our highlights cover accomplishments for the first 6 months of work.

The research has three objectives: 1) to advance computer science research by developing an expressive web services description language and techniques for dynamically composing web services, 2) to develop and conduct test applications of an intra-metropolitan goods movement flow model using web services in cooperation with government partners, and 3) to use the model to conduct social science research on intra-metropolitan economic linkages and spatial structure. Although the focus is on the specific topic of urban goods movement, the approach to web service composition is general and can be applied to other scientific data gathering and analysis tasks.

The first objective is to develop Argos, a general framework for dynamically composing web services. Many scientific problems can be modeled as a workflow that includes information gathering and processing operations. We propose a unifying framework where these operations are modeled as web services and the scientific workflows as compositions of web services. Argos will provide graphical tools for manual specification and composition of web services, as well as automatic composition based on expressive web service descriptions for given application domains (such as transportation planning). The second objective is to use Argos in an actual metropolitan planning application. In consultation with an advisory team of government representatives, a scenario analysis using the Los Angeles region as a case study will be conducted. The third objective is to extend the transportation planning domain to address problems of urban spatial structure that heretofore have not been practical for social science researchers to study due to the lack of tools for integrating and analyzing available data.

Our Year 1 work has focused on developing the Argos framework. We are developing techniques to automatically create computational workflows in response to user data requests. We propose to integrate data from secondary sources to estimate the flow of commodities in the Los Angeles region, following Gordon and Pan [2001]. The Argos approach presents several advantages: 1) improved cost-effectiveness due to automated data integration and processing; 2) flexibility in using and combining data sources and operations; 3) ease of updating by using data from live sources whenever possible.

We have developed our initial design for automatic composition of web services for goods movement analysis and are presenting it as a demo at this conference. The work consists of three steps:

1. Define an ontology of the domain to model data processed throughout a workflow.
2. Describe data sources and operations using this ontology.
3. Automatically compose a workflow in response to a user request based on mediator techniques.

First, we have developed an ontology of the goods movement domain to describe the data provided by the sources and the data utilized as input and output of the processing operations (see Figure 1). We consider each data item as a measurement that has values along a set of dimensions, such as geographic area (e.g., Long Beach), type of flow (e.g., imports), type of product (e.g., vehicles and parts), time interval (e.g., January 2001), value and unit (e.g., 1108 thousand short tons). Some of the values of the dimensions have a hierarchical structure, either *type* or *part-of* hierarchies. To facilitate the knowledge

---

[1] Ambite, Harth and Decker: Information Sciences Institute, University of Southern California
[2] Giuliano, Gordon, Pan, Wang: School of Policy, Planning and Development, University of California

acquisition task we used the Protégé ontology editor [Noy et al., 2001] (see Figure 1). As a common syntax for the ontology and the data we adopt the Resource Description Framework (RDF) [Lassila & Swick, 1999]. RDF represents labeled graphs (i.e., semantic networks). RDF also includes edge labels with predefined class/subclass semantics.

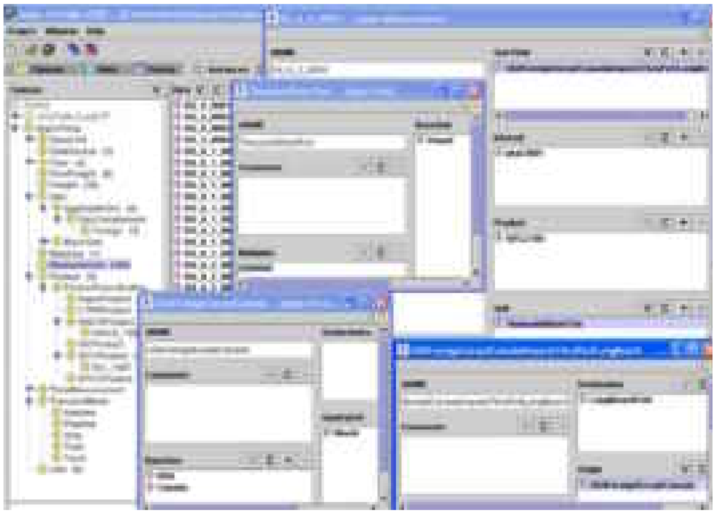

Figure 1. Ontology for Goods Movement Analysis

Second, we have modeled the data sources and operations uniformly as web services that produce or transform measurements. To describe the semantics of these web services, we use the Triple language [Sintek & Decker, 2002]. We used Triple logic rules built with terms from the ontology to accurately specify the inputs and outputs of sources and operations.

Third, we are developing an approach to web service composition based on mediator techniques [Thakkar et al., 2003]. Given a user data request, our system uses the Triple service descriptions (sources and operations) to generate a logic program that answers the request. What differs our work from [Thakkar et al., 2003] is that we used a richer representation language with object-oriented features as opposed to their flat relational model. Moreover, we model more complex transformations, not just relational algebra operations. In fact, the operations can be arbitrary algorithms or Triple programs.

We have taken a subset of the sources and operations of our commodity workflow and developed a working prototype. We implemented these sources and operations as web services that exchange RDF data, and created a manual service composition based on BPEL4WS [2003] that can be queried both for intermediate and final data. Our next steps include the generation of automatic web service composition in response to user requests, refining our approach, and expanding our web service composition techniques to the more complex parts of the workflow.

References

Ambite, J.L., Giuliano, G., Gordon, P., Harth, A., Wang L., Pan, Q., and Decker, S., (2004) "Argos: An Ontology and Web Service Composition Infrastructure for Goods Movement Analysis." System Demo, dg.o2004.

BPEL4WS (2003), Business Process Execution Language for Web Services Version 1.1, 05 May 2003, http://www-106.ibm.com/developerworks/library/ws-bpel/

Gordon, P, Pan, Q. (2001) "Assembling and Processing Freight Shipment Data: Developing a GIS-Based Origin-Destination Matrix for Southern California Freight Flows." Final Report, Project METRANS National Center for Metropolitan Transportation Research. (www.metrans.org).

Lassila, O, Swick, R. (Eds) (1999). Resource Description Framework (RDF) Model and Syntax Specification, *W3C Recommendation* 22 February. 1999. http://www.w3.org/TR/RECrdf-syntax/

Noy, N.F., Sintek, M., Decker, S., Crubézy, M., Fergerson, R., Musen, M.A (2001). Creating Semantic Web Contents with Protégé-2000. In *IEEE Intelligent Systems*, 16 (2), 60-71. Special Issue on the Semantic Web. March/April 2001.

Sintek, M., Decker, S (2002). TRIPLE - A Query, Inference, and Transformation Language for the Semantic Web. *International Semantic Web Conference (ISWC*), Sardinia, Italy, June 2002.

Thakkar, S, Knoblock, C. A., Ambite, J. L. (2003). A View Integration Approach to Dynamic Composition of Web Services *Workshop on Planning for Web Services*, The 13th International Conference on Automated Planning & Scheduling (ICAPS 2003), Trento, Italy, 2003.