Position paper for Specialist Meeting on Spatial Webs, NCGIA, December 2-4, 2004

Developing Intelligent Software Agents for Semantic Interoperability on Internet GIS.

Author: Ming-Hsiang (Ming) Tsou, Department of Geography, San Diego State University

The progress of geospatial information technology and the need for global distribution of geographic information is pushing the GIS community to develop Geographic Information Services (GIServices) on the Internet. Currently, on-line GIServices are well adopted in academic research projects, federal and local governments, and the GIS industry. Data sharing is now commonplace, but the lack of semantic interoperability becomes a major hurdle to advance the development of distributed GIServices. This paper describes a software agent-based communication mechanism (GeoAgents) that dynamically integrates geospatial data and GIS services across the Internet. The goal of Internet GeoAgents is to reduce user work and information overload by interpreting, filtering, and converting semantic information automatically. The adoption of intelligent software agent technology can facilitate the semantic interoperability on Internet GIS applications. Three fundamental roles for GeoAgents are essential to distributing GIServices: information broker, information interpreter, and decision maker.

An information broker helps users search requested information, and filter out unnecessary elements according to user-supplied specification. The interaction process between software agents and users is similar to the relationship between travel agents and airplane travelers. When a traveler tells the travel agent their preference, i.e. their dates of travel and the names of departure and arrival cities (a task specification), the travel agent filters and simplifies information from complicated flight schedules to create a few possible itineraries for the traveler to choose from. The travel agent will in theory only provide a reasonable range of choices to the buyer. Similarly, a software agent (information broker) will filter out unnecessary information and provide a reasonable number of choices to the user.

An information interpreter agent can access and convey information. In distributed network environments, heterogeneous data models and systems cannot communicate directly. Geospatial data present special challenges as they are characterized by spatial dependence, spatial heterogeneity, and sensitivities to spatial and temporal resolution. These properties complicate data fusion and transfer protocols. A GeoAgent can bridges heterogeneous information islands to convert between different data types and models. To translate the information correctly, a GeoAgent must acquire related GIS ontology and semantic translation methods. The knowledge and methods can be previously defined and encapsulated in the metadata of object-based components or be provided by external ontology databases/entities.

A decision maker agent collects and analyzes information according to specific events, makes an optimal decision in collaboration with a user or other agents, and executes a behavior accordingly. The behavior of decision maker agents is based on the rules defined by its own knowledge base, by user-defined rules, or by the suggestions from other software agents. Decision maker agents require appropriate user interfaces for accepting decision-making rules and for procedure verification. Also, appropriate communication protocols must be formalized to enable message passing for agent interaction. Another issue involves how an agent chooses

collaborative agents for the decision making process. Since a decision maker agent will have more power than the other types of agents, a hierarchy of software agents must be defined. For example, in a network security system, a certificate authority (CA) server (higher level) can assign or distribute public keys or private keys to the client web servers (lower level) for improving the security of web browsing and transactions. To obtain the certificates, the client web servers need to send their requests to the CAs. The CAs (decision makers) then decides whether to issue the certificates to the client servers or not.

The advantage of adopting software agents is that complicated GIS tasks can be executed under software agent control to provide more comprehensive services for information access, retrieval and processing. Consider the following request: *Finding GIS data [Roads] in [San Diego]*. In a current Web search engine process, the search can only have two possible answers: Found, or Not-found. In an intelligent software agent framework, software agents can create more different answers as the following:

Request *Software Agent Response*
(Find, [Roads, San Diego]) → (Please re-define the keyword "San Diego" for 1: City 2. County)
→ (Search Result A: Location = Web site: GIS.SDSU.EDU)
→ (Alternative choice: replacing [Roads] by [Major Roads], then search)
→ (Alternative choice: selection [Roads, California] from Web-B with additional GIS process required (Clipping [Boundary-of-San-Diego] from California)

One important criterion for the implementation of intelligent software agent framework is the establishment of GIS ontology. An ontology is a particular conceptualization of a set of objects, concepts, and other entities about which knowledge is expressed, and of the relationships among them. GIS ontology defines knowledge that guides GeoAgents' behavior. The following statements provide a few example statements of GIS ontology.

- *The map units in State Plane Coordinate System are feet.*
- *1 feet = 0.3048 meters.*
- *'Roads' include 'major roads, highways, unpaved roads'.*

By adopting these rules, a software agent can make inferences to accomplish tasks. For example, if a geodata agent receives a request for "buffer at 200 feet" on an ArcView Shapefile with UTM coordinate system (meter-based). The geodata can retrieve these rules from the ontology and convert 200 feet to 60.96 meters, and revise the buffer operation accordingly. To demonstrate the feasibility of GeoAgents, a Java-based prototype of an intelligent data processing system for environmental models was developed to help end users search and retrieve geographic information from distributed and heterogeneous data sources via the Internet. The adoption of GeoAgents in the prototype can automatically convert the retrieved data to readable formats required by specific environmental models.

To summarize, "sharing" is the most essential motivation of distributed GIServices. To sharing data, information, and knowledge among the GIS community, we need to develop a flexible framework to exchange semantic information across networks. Adopting intelligent software agents has great potential to facilitate such framework for semantic interoperability.