

Information Agents

Jane Hsu

Information Gathering Problem

Growing number of information sources
Distributed heterogeneous information sources

- Traditional databases
- Flat files
- Knowledge bases
- Programs etc.

Traditional approaches to locating, combining, processing, & organizing info do not scale well.

Information Gathering Systems

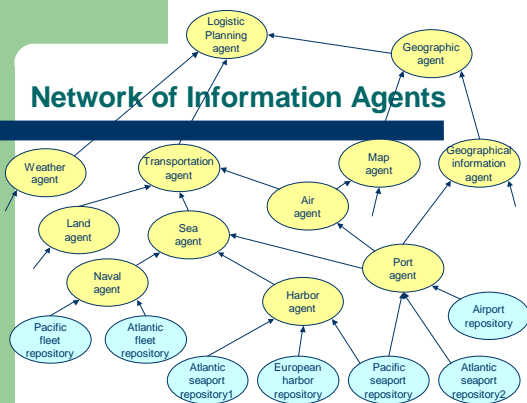
SIMS [Arens, Knoblock et al, 1993]
The Ahoy! Softbot [Shakes et al, 1994]
Occam [Kwok & Weld, 1996]
InfoSleuth Agents [MCC, 1996]
Information Manifold [Levy et al, 1995]
InfoMaster [Geddis et al, 1995]
Tsimmis [Stanford, 1994]
XML View Inference [Jeong, 2001]

SIMS [USC/ISI]

A network (DAG) of information agents
Each agent provides expertise on a specific topic by drawing on relevant information from other information agents

- Domain model, i.e. ontology
 - Information source model
- Loom knowledge representation language
KQML communication language and protocol
Information mediator

Network of Information Agents



Query Processing

Develop a plan for obtaining requested data

- Select the information source
- Processing operations
- Sites where operations will be performed
- The order in which to perform them

Desirable features

- Parallel execution
- Augment and re-plan queries that fail
- Gather information to aid query processing

Learning in SIMS

Cache frequently retrieved or difficult-to-retrieve information
Learn about the contents of the information
Semantic query processing

Features

Modularity
Extensibility
Flexibility
Efficiency
Adaptability

U1760 Term Project

System design: 11/7
Agent design: 11/14
Mid-term report: 11/28

Task Domain: Technical Reference

Scenario One:

Given: query expression
Find: relevant bibliographical data
; [The Collection of CS Bibliographies](#)
; [DBLP: Computer Science Bibliography](#)
; [Networked Computer Science Technical Reference Library](#)

Scenario Two:

Given: bibliographical data
Retrieve: the target paper
; BibAgent

Tools

Online
; Digital library or online repository
; Library catalogue
; Search engine, e.g. Google
; Publisher, e.g. IEEE, ACM, AAAI, universities
; Bookstore, e.g. Amazon, Barnes&Noble
; Specialized catalogue
; Personal homepage
Offline
; Library
; Bookstore

Problems Encountered

Incomplete citation
; Only title, keywords, or author
Incorrect citation
Non-existent document
Inaccessible information sources
Multiple information sources
Non-standard document format
Implicit information

