Multi Agent System and an Approach for Association Rule Mining

Imran Qureshi
Department of Information Technology
Guru Nanak Institute of Technology, Ibrahimpatnam, Telangana, India

K Suresh
Department of Computer Science Engineering
Guru Nanak Institute of Technology, Ibrahimpatnam, Telangana, India

Rama Murthy Garimella
International Institute of Information Technology, Hyderabad, Telangana, India

Abstract - This work addresses an architecture that not only allows mining data sources of multiple formats, but also allows extensibility for future formats. It also presents a methodology for unification of association rules prior to incorporation in a BKB. It will use an existing agent development tool to establish a multi-agent based framework and define the communications between those agents. The framework will be designed to accept a request from PESKI for one of three possible data mining operations. These operations are discussed in more detail in Chapter 4. Once the system accepts the request, it will determine which data sources can fulfill the request and tasks the agents responsible for those sources to begin data mining. Once results have been obtained, they will be unified to eliminate redundant or conflicting results and returned to PESKI. Two new data source formats will be introduced into the PESKI schema and they will be mined for association rules, the results unified into a unique list of results, and then passed back to PESKI for incorporation into the BKB. The process will be automated and will use existing message passing formats to communicate with PESKI.

I. INTRODUCTION

An Agent is a virtual individual or object that follows rules written in code. An agent is like a puppet, it acts according to our movements. For example agent modeling for ecosystem should have rabbit and grass as agents. These agents have some characteristics. The characteristics for rabbit and grass will be name, type, color and motion (not limited to this, can have any characteristic according to type of modeling, these characteristics have been just taken as an example). In the former case name is grass type is plant color is green and motion is none. In later case name is bunny (just for example can be any name user wants to keep) type is animal color is green and motion is walk. An agent may also have behavior. For example rabbit should have a behavior like movement. In computer science agent's behavior are procedures or methods we write. Agents are able to interact with each other. A computer agent can be made intelligent by writing pieces of software codes

II. FOUR RULES

The four rules an agent should have are
Rule 1: Ability to perceive environment (A robot will have sensors and human eye have eyes to perceive)
Rule 2: Observations used to make decisions
Rule 3: Decisions will result in action (Like do something or respond)
Rule 4: Any action taken must be rational i.e. best possible action.
In order to make intelligent this rule is very important. It has been proved that humans are not good at being rational because of our emotions and beliefs

III. CATEGORIES OF AGENTS IN PROPOSED SYSTEM

The seven main categories of agents in the proposed system are

1) User agent
2) Task agent
3) Broker agent
4) Ontology agent
5) Data analysis agent
6) Unification agent
7) Registration agent

Interconnection and Distribution has become core topics in Computer Science. But the Interconnection and the Distribution are coupled with the need for systems to represent our best interests that is the systems that can *cooperate* and reach agreement with other systems that have different interests (much as we do with other people). The above issues were not studied in Computer Science until some time ago. All of these trends have led to the new emerging field called multi agent systems in computer science

IV. DEFINITION OF AN AGENT

Agent is a computer system or a program that is capable of independent action on behalf of its user or owner

Agents interact with each other to perform a specific. Agents need to cooperate and coordinate and in some cases agents need to negotiate with each other to complete the given task

Number of questions arises while building a Multi agent system such as kind of language used to communicate between agents, how conflicts are recognized and how to reach an agreement. How Autonomous agents achieve their goals. Multi agent systems are unique because agents can compute and process the information too. Each researcher has his own view about MULTI AGENT SYSTEM but the researchers have much in common. An example of MULTI AGENT SYSTEM is searching a answer for a specific question in internet is difficult like to get information about different flights. Instead of searching the details by visiting respective sites why can't an agent do it on behalf of user? The agent is given a query to retrieve the desired information, the agent visits the different sites and displays the result to the use (Example is searching the flight details in makemytrip.com). Failure occurs if information is unavailable or because of network failure

Different Research issues arise like what parameters an agent require, what algorithm is used by agent to perform the task etc. MULTI AGENT SYSTEM is influenced by many fields like game theory, philosophy, ecology, social sciences etc. There are many criticisms for multi agent system also some say that isn't is a just concurrent or distributed system. The answer is MULTI AGENT SYSTEM is not same as Distributed system or AI because agents in MULTI AGENT SYSTEM are autonomous and are capable of taking decisions. Examples of non-trivial agents are thermostat and daemons in UNIX. The main property of intelligent agent is flexible autonomous action. Flexible means an agent should be reactive, proactive and social

A reactive system responds to environment change an agent decides whether this action is worth taking or what to do in case of failure. Proactiveness refers for achieving the goal i.e. goal directed behavior.

Agents should be able to interact with other agents using any communication language like KQML and cooperate with each other. An agent should be able move around a network called mobile agents and veracity property of agent refers that the agent knowingly does not give the false information. The other properties of agent include benevolence (agents do not have conflicting goals), rationality and learning

Architecture of an agent can be any one of the following three architectures

1) Symbolic/logical
2) Reactive
3) Hybrid

More about above three architectures can be read from book written by Michael wool ridge title an introduction to Multi Agent System

Communication of agents can be done using KQML (Knowledge query and manipulation language) & KIF to cooperate with each other. Speech act is the inspiration for communication among Multi agent system. Two general components of speech act are per formative i.e. request/inform/promise etc and the second one is propositional content (for example the door is closed). An example for speech act is

Per formative= inform

Content=”holiday”

Speech act = “Tomorrow/today is a holiday”

Defining the semantics of speech act is important because when someone has uttered something how do the agent know whether it is a request or inform. More on semantics can be studied in the research of Cohen and Perrault in which they defined the semantics of speech act.

Now we look at Agent communication language (ACL). The best known ACL is KQML which is developed by ARPA. It has two parts one is knowledge query manipulation language and Knowledge interchange format. KQML defines various per formatives like perform, tell, reply etc. KIF is a language for message content which state the relationship (A is B’s Boss), properties of things etc. Ontology refers to common terms or specification agreed by agents to communicate

An for Example KQML/KIF dialogue is given below
Alice to Bob: (ask-if (> (size object1) (size object2)))
Bob to Alice: (reply true)
Bob to Alice: (inform (= (size object1) 30))
Bob to Alice: (inform (= (size object2) 28))

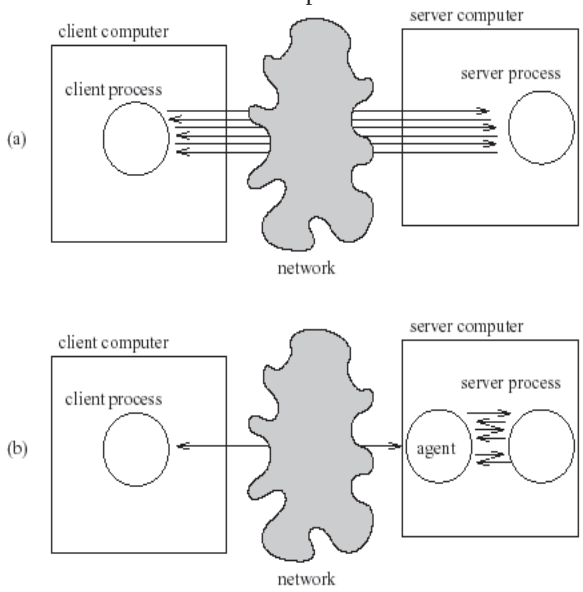
FIPA recently started work on agent’s standard. The basic structure of FIPA is similar to KQML

A wide range of per formatives defined by FIPA are given below

performative	passing info	requesting info	negotiation	performing actions	error handling
accept-proposal			x		
agree				x	
cancel		x		x	
cfp			x		
confirm	x				
disconfirm	x				
failure					x
inform	x				
inform-if	x				
inform-ref	x				
not-understood					x
propose			x		
query-if		x			
query-ref		x			
refuse				x	
reject-proposal			x		
request				x	
request-when				x	
request-whenever				x	
subscribe		x			

Inform and Request are two basic per formatives in FIPA. The inform and request is defined into two parts they are one pre-condition (what must be true for speech act to succeed) and second one is rational effect (what the sender of message hopes to bring out)

Difference between Remote procedure call and mobile agents



Need for mobile agents are because it takes low-band-width thus making an efficient use of network resources. The main application of MULTI AGENT SYSTEM are developing distributed systems over a network

Association rule mining is a data mining technique used to extract useful trends from huge amount of data base. Data can be retrieved from many sources like flat files, Relational data bases and other information sources. As the different data sources are used to extract information the problem of how to integrate/combine these heterogeneous data sources. And also within a single business organization data is separated over different geographic locations with varying formats. Thus we have huge amount of data for data mining association rules. Thus the main aim of this research is to perform association rule from multiple heterogeneous data sources using multi agent system and then unifying the result to knowledge base Data mining is the process of extracting useful information from data which is unstructured. Data Mining is also called Knowledge data discovery or data dredging. The different stages of Data mining include preprocessing, data reduction used to eliminate the variables which is of no Interest. Then the next step is to select the data mining task and execute. The results obtained should be presented to the user in understandable format Association rule is used to find the relationship between two values of the form a→b. That is how likely a customer purchases b when he purchases a. Association rule mining is used in market Basket analysis. Two measures of interest are support and confidence. Only those rules are taken into consideration which satisfies minimum support and confidence. Support is set by the user. Support should be selected in such a way that it is not too low (many rules are generated) and not Too high (No or very few rules are selected)

V. PROBLEM ANALYSIS

This framework will utilize a current specialist's improvement device to build up a multi-agent based structure and characterize the interchanges between these agents. The structure will acknowledge a solicitation for one of three conceivable information mining operations. Once the framework acknowledges the solicitation, it will figure out which information sources can satisfy the solicitation and undertakings the operators in charge of those sources to start information mining. When results have been gotten, they will be brought together to kill excess or clashing results.

Examination of the issue demonstrates that there are two primary area level ideas that must be used – interface and preparing. The new framework must interface with an outside application to get the information mining tasking. When it gets the tasking, it must process the information and decide the correct information sources. Information sources are then mined (still under handling) and results are brought together, and then displayed back to the application (interface).

VI. AGENTS FRAME WORK

In light of the issue and environment investigation, there are a few sorts of specialists in this framework. They are assembled in light of the similitude of the errands they perform and their individual objectives. The seven fundamental classes of specialists in this framework are User, Task, Broker, Ontology, Data Analysis, Unification and Registration. Every falls inside of the space level ideas determined in the issue and environment examination stage. The User, Task, Broker, Ontology, and Registration Agents are all interface specialists. They all give interfaces to either an outside framework, specialists inside of the framework, or information sources. They don't handle the information in any capacity. The Data Analysis and Unification Agents are handling operators which control the solicitation or information inside of the framework.

VII. ASSOCIATION RULE MINING USING MULTI AGENT SYSTEM

Keeping in mind the end goal to see how the procedure functions, it is helpful to follow through the correspondence ways a solicitation/request would take. The procedure starts with the User Agent accepting notice from the application that a solicitation needs handling. The User Agent gets the application-designed demand and changes over the information into a solicitation of sort Request Class so that the specialists framework can comprehend it. It then makes an impression on the Task Agent. The Task Agent then approaches the Broker Agent for all valuable Data Analysis Agents. The Broker Agent gets the solicitation, incorporates a rundown of all Data Analysis Agents in the framework, and then verifies whether an Ontology Agent exists. On the off chance that one exists, it sends the rundown of Data Analysis Agents, alongside the solicitation for examination. The Ontology Agent acknowledges the solicitation and checks to guarantee the Data Analysis spaces against the solicitation. It gives back a rundown of helpful operators to the Broker. The Broker then returns this rundown to the Task Agent. When the Task Agent gets the rundown, it sends a solicitation to every Data Analysis in the rundown to start mining. Every Data Analysis Agent acknowledges the solicitation and starts information mining their appropriate information source. Once finished, the Data Analysis Agents sends the outcomes back to the Task Agent. When the Task Agent has every one of the outcomes, it passes them to the Unification Agent. The Unification Agent forms the outcomes, binds together them, and passes the outcomes back to the Task Agent. The Task Agent then passes the brought together results back to the User Agent. The User Agent then changes over the outcomes to an arrangement the outside application can perceive and informs the application the outcomes are accessible.

BIBLIOGRAPHY

[AGRA96] Agrawal, Rakesh, et. al. “Fast Discovery of Association Rules” *Advances in Knowledge Discovery and Data Mining*. AAAI/MIT Press. Chapter 12, 307-328.

[ACHK93] Arens, Yigal, Chin Y. Chee, Chun-Nan Hsu, and Craig A. Knoblock *Retrieving and Integrating Data from Multiple Information Sources.* International Journal of Intelligent and Cooperative Information Systems. Vol. 2, No. 2. Pp. 127-158, 1993.

[AS94] Agrawal, Rakesh and Ramakrishnan Srikant “Fast Algorithms for Mining Association Rules.” September 1994 *Proceedings of the 20th Very Large Data Bases Conference*, Santiago, Chile 487 – 499.

[BAY96] Bayardo, R. et. al. “Semantic Integration of Information in Open and Dynamic Environments” MCC Technical Report, MCC-INSL-088-96, 1996.

[BRAD97] Bradshaw, Jeffrey “Software Agents” J. Bradshaw ed., AAAI/MIT Press, Menlo Park, CA 1997.

[CHAU97] Chauhan, Deepika “JAFMAS: A Java-based Agent Framework for Multiagent Systems Development and Implementation”, ECECS Department, University of Cincinnati, 1997

[FU95] Usama M. Fayyad, Ramasamy Uthurusamy (Eds.): August 20-21, 1995 Proceedings of the First International Conference on Knowledge Discovery and Data Mining (KDD-95), Montreal, Canada, AAAI Press.

[JAF97] JAFMAS, Comparison of JAFMAS to other Java-Based Agent tools, http://www.ececs.uc.edu/~abaker/JAFMAS/tab2.html, University of Cincinnati, 1997.

[JAT97] JATLite, JATLite Overview, http://java.stanford.edu/java_agent/html/, Stanford University, 1997.

[MHIP95] H. Garcia-Molina, J. Hammer, K. Ireland, Y. Papakonstantinou, J. Ullman, and Jennifer Widom. "Integrating and Accessing Heterogeneous Information Sources in TSIMMIS". In *Proceedings of the AAAI Symposium on Information Gathering*, pp. 61-64, Stanford, California, March 1995. .

[HNFD98] J. J. Han, R. T. Ng, Y. Fu, and S. Dao, `` Dealing with Semantic Heterogeneity by Generalization-Based Data Mining Techniques", M. P. Papazoglou and G. Schlageter (eds.), *Cooperative Information Systems: Current Trends & Directions* , Academic Press, 1998, pp. 207-231.