A System for Multi-agent Information Retrieval

Paula Mc Dermott and Colm O'Riordan

Department of IT, NUI, Galway
{Paula.Mc_Dermott,colmor}@geminga.nuigalway.ie

Abstract. In this paper we discuss a multi-agent information trading system, which we argue provides a suitable and flexible design for distributed information retrieval, and a suitable test-bed for studying agent cooperation, coordination and negotiation. Within the fields of Information Retrieval (IR) and Information Filtering (IF), agents represent a suitable paradigm to conceptualise, design and implement an IR system.

1 Introduction

The purpose of the system described and developed in this paper is twofold. Firstly, the system acts as an agent based information trading system and secondly it forms a framework for testing various cooperation, coordination and coalition formation strategies. Within this framework, agents may have the potential to learn about information needs and other agents abilities etc. As sources become more distributed, we have a set of information providers, with their set of repositories and an associated cost to access this information.

The paper is laid out as follows. Section 2 outlines research in IR and distributed IR. Research and development in MAS is outlined in Sect. 3. The following section discusses the design of our system. Section 5 describes current and future experiments. Section 6 offers results and the paper is concluded in Sect. 7.

2 Information Retrieval

The task of an Information Retrieval system is to retrieve documents containing information that is relevant to a user's information need [1]. This task is a difficult one as it usually deals with natural language which is not always well structured and may be semantically ambiguous [2].

2.1 Approaches

The task of IR can be broken down into three subtasks, namely, representation, comparison and feedback.

Documents and queries must be represented in the same machine readable manner, such that the system can perform comparison techniques between them.

M. O'Neill et al. (Eds.): AICS 2002, LNAI 2464, pp. 70–77, 2002.

Representation techniques range from using indexes, vector and matrix representation to more modern representations, such as neural networks, connectionist networks and semantic networks [3].

The comparison mechanisms used are dependent on the underlying representations employed in the system. There are three classic representations (models) in IR, namely, Boolean [2], Vector [6] and Probabilistic [19].

Due to the nature of IR the initial results will be a partial match to the query. The returned set is improved by query reformulation, which may be performed by the user but is largely automated by IR systems. It involves two steps: expanding the original query with new terms and re-weighting the terms in the expanded query.

2.2 Distributed Information Retrieval

Given the increase in size of information repositories, the need for alternative architectures and algorithms has emerged. Modern information environments have become large, open and heterogeneous [7]. This has led to the increased adoption of both distributed and parallel architectures.

Commonly identified problems within distributed IR are those of source selection (which repository is most appropriate for a given query) and result merging (how to combine results from different sites). Possible solutions for source selection include central indexing mechanisms and rule-based techniques. Voorhees et al. [18] and Callan et al. [17] provide techniques for dealing with result merging.

3 Multi-agent Systems

Software agents date back to the early days of AI work and Carl Hewitt's concurrent actor model [5]. In this model, Hewitt proposed the concept of a self-contained interactive and concurrently executing object which he termed an actor. This object had some encapsulated internal state and could respond to messages from similar objects [4]. Wooldridge [16] provides the following definition: *"An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives"*.

Generally a multi-agent system is composed of a number of agents that are able to interact with each other and the environment and that differ from each other in their skills and their knowledge about the environment [8]. The agents in a multi-agent system will need to communicate with each other, be aware of each other and reason about each other. The complexity introduced by agent-agent interactions has created, among others, the research areas of agent cooperation, coordination and coalition formation.

3.1 Agent Traits

Cooperation. Cooperation is often presented as one of the key concepts which differentiates multi-agent systems from other related disciplines such as distributed computing, object-oriented systems and expert systems [9]. Notions of shared or joint goals and notions of complimentary individual action have been posited. Theories from formal logics [10], philosophy [15] and game theory [14] have been adopted to reason about cooperation and to design and develop models whereby agents cooperate.

Coordination. Coordination is a central issue in software agent systems and in Distributed Artificial Intelligence (DAI) [11]. Coordination is required in a multi-agent system for many reasons, including: preventing anarchy or chaos, meeting global constraints, efficiency, distributed expertise and dependencies between agents' actions. It is important to note, that coordination may require cooperation but cooperation will not necessarily result in coordination. Many coordination strategies have been developed, they can be categorised as follows: Organisational Structuring, Contracting, Multi-Agent Planning and Negotiation [11].

Coalition Formation. A coalition can be defined as a group of agents who have decided to cooperate in order to perform a common task [12]. Shehory, Sycara and Jha [12] believe that the incorporation of a coalition formation mechanism will increase the efficiency of group-wise task execution, resulting in near-optimal task performance. In general agents will only form a coalition if each member of the coalition gains more by joining the coalition than by working alone. Work in this domain ranges from sub-additive coalitions [13] to greedy strategy coalitions [12] and those that are based on agent negotiation.

In recent years, agents have been developed and 'released' into information management environments in an attempt to improve the mechanisms for finding, fusing, using, presenting, managing and updating information.

4 System Design

The primary concern of this paper is the application of agent-based systems to the domain of information management. Weiss [8] describes information agents as: agents that have access to multiple, potentially heterogeneous and geographically distributed information sources. Information agents have to cope with the increasing complexity of modern information environments and retrieve, analyse, manipulate and integrate information from different sources.

4.1 Agent Community

The agents in the system have three main components in their architecture: *knowledge, beliefs and capabilities.*

The agent's knowledge base comprises *user(s)*, *information need(s)* and a *utility* value. The information need represents the task that the user is requesting the agent to carry out. The utility or self-belief value is a measure of the user's satisfaction with the agent's performance. Depending on the agent's performance the user will increase, or decrease, its trust in the agent through user feedback.

The agent's beliefs represent the agent's opinions of one another. The *belief* rating refers to one agent's level of belief in another agents ability to perform a task. The *trust* rating, on the other hand, relates to an agent's belief that the agent will actually carry out the task. This section is an important part of the agent architecture. The belief and trust levels that an agent has in another agent will be taken into account when an agent is considering who is most suitable for a particular task.

The capabilities section refers to the tasks that the agents will be able to perform. All agents will be capable of carrying out the same tasks (including *retrieval,* and *learning*) but they will differ in their ability to perform them, for example some agents will perform Boolean retrievals while others will perform vector based retrievals.

4.2 Interactions

The system is relatively straight-forward if the agents are capable of performing the task specified by their users. They simply query their own repositories, using their IR module, and return the retrieved documents to their user. This acts as a set of separate centralised IR systems.

The more complicated and interesting aspects of the system can be seen if an agent is unable to perform the task required (due to poor capabilities or repositories). If, for example, the agent in question does not have access to the information needed to fulfill its task(s) then it must employ the help of another agent. This is quite a complex task. In such a scenario agents must be able to communicate with one another, propose, reject, accept and counter propose courses of action. This situation may also occur on a larger scale. It is possible for an agent to 'employ' several agents to perform the task in question (coalition formation).

When an agent performs a task it expects a reward (modelled as an increase in user's trust in an agent). When a group of agents work together on a task they must also be rewarded for their efforts. This introduces the notion of a payoff scheme. The bonus available for the task is still the same regardless of how many agents work on the task. The bonus must be split in some way between the agents. Several coalitions may form within one agent community. There can be coalitions within coalitions. It should be noted that the agents are under no obligation to join in any coalition or even to cooperate with other agents.

In summary, we have a set of repositories, agents and users. These can be easily expanded if required. Primitives and protocols exist which allow retrieval, trading and coordination schemes (coalition formation). We envisage that this architecture can be re-used with different mechanisms in place in order to compare approaches.

5 Experimental Setup

5.1 Information Retrieval Systems

Information collections used in the system are standard IR test collections. A collection of abstracts from the Communications of the Association of Computing Machinery (CACM), the CISI collection and the Cranfield collection provide a substantial amount of documents for use in the system. Each agent in the system has the ability to search its own document collection. We require that different agents have different capabilities so we have several IR engines. The IR modules currently available in the system are: the Boolean model and the Vector Space model.

5.2 Multi-agent System Protocols

The main motivation behind the test-bed implementation is to allow various agent cooperation, coordination and coalition formation strategies to be easily examined. We hope to measure complexity (messages passed), fitness and user satisfaction.

An agent may want to form a coalition with another agent or set of agents if it is unable to perform the task required of it. In order to form a coalition the agent has to carry out the following steps:

- Select agent(s) to join the coalition based on belief and trust ratings obtained from previous interactions.
- Broadcast *propose* performatives to selected agent(s).
- Await responses (*accept, reject or counter-propose* performatives.)
- If the accepting agents are happy with the terms of the coalition then the contract is implemented. On the other hand if the agents are not satisfied with the terms of the contract then they will try to re-negotiate with the offering agent by counter-proposing. This cycle may have several iterations until the agents decide to join or to finally reject the coalition
- The contract is implemented. Each participating agent performs its given task.
- Finally the initiating agent modifies its belief and trust ratings of the participating agents based on their performance in the coalition.

Variations on the above can easily be added.

5.3 Experiments

Initial experiments involve the simulation of a distributed multi-user IR system wherein users information needs are satisfied via a set of information agents. A number of information repositories will be created and will be accessible by some or all of the agents. The goal again was twofold: firstly to conduct experiments regarding agents in an IR domain, and secondly to demonstrate its suitability as a framework for testing agent protocols.

Initial agent experiments will involve local updates of belief and trust models possessed by individual agents, experiments regarding coalitions and normative constraints can be easily incorporated into the existing system. The system, as it is designed, allows users, n agents, n information needs and m IR modules to interact within the agent community. The n agents satisfy the users by providing information, through the use of m different IR modules.

6 Results

6.1 Experiment 1 – Self-Belief

The first experiment involves monitoring changes in the agent's confidence over time. This is indicated by user satisfaction. This value is modified through simulated user feedback (this is achieved through the use of human relevance judgements which are provided with the collections) and changes after each retrieval. Initially the agents are assigned the value 0.5 (on a scale 0-1), to represent their beliefs. This suggests that they have neither a strong nor weak belief in their own ability. There are ten agents in the system, half of which perform vector-based retrievals, with the remainder performing Boolean retrievals. Figure 1 illustrates the average belief for the vector and Boolean agents after each retrieval is performed. Each agent performs twenty queries which are randomly selected.

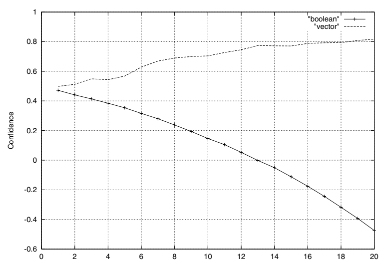

Fig. 1. Average performance of Boolean and vector agents

Figure 1 illustrates that the average user's trust in a vector-based agent is significantly higher than that of a Boolean agent. This indicates that the user has a higher confidence in, and is more satisfied with, vector-based agents. This allows us to posit that a Boolean agent will need to interact on a more frequent basis with other agents, preferably with vector agents, in order to satisfy its user.

6.2 Experiment 2 – Belief in Others

This experiment allows the agents to function as before but with the addition of the ability to modify their belief in other agent's abilities. This experiment allows us to show that, over time, an agent can learn to trust or mistrust another agent. If an agent is unable to meet its user's information need, or it has experienced a decrease in its user's trust, it makes an attempt to 'employ' another agent to aid with a task. In order to select the agent to learn from, the agent examines its belief ratings. Initially these are set to 0.5, and the agent makes its choice randomly. The agent communicates with the selected agent, and asks it to perform the task. Again the selected agent will have no strong belief in any other agent, and it randomly decides whether or to accept or reject. Upon acceptance of the offer the selected agent searches its database and returns the results set to the 'employer' agent, who then returns the documents to the user as its own work. A successful interaction results in an increase in the agent's utility and its belief in the contracted agent.

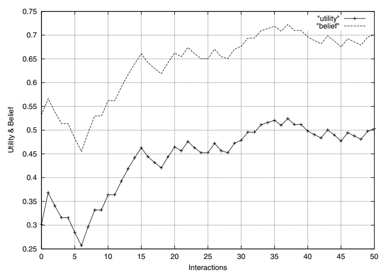

Fig. 2. Performance and belief over time

Figure 2 represents sample changes in self-belief (utility) and belief in other agents employed over time. As we can see this selected agent performs quite well over the first two queries and then deteriorates. When the 'employer' agent notices its belief in this agent decreasing, it selects another agent. Immediately we see a steady increase occur in both the utility and belief values. Upon examination of the agents involved we notice that the initial agent performs Boolean retrievals while the second agent is vector-based. We can see that agents learn over time to have greater trust in more reliable agents.

7 Summary and Conclusion

The areas of IR and multi-agent systems represent different areas of study but can be combined in order to provide a robust, extendible, intelligent, potentially distributed information management system. We believe that this approach provides a unique method of dealing with the traditional IR problem. The agents provide an intelligent module, in which they have the ability to communicate with, learn from, trust and distrust each other. We believe that this process of information retrieval provides a better quality of service to the end user. The multi-agent paradigm undoubtedly extends the capabilities of IR modules in a distributed environment. We also posit that the system developed acts as a useful test-bed for experimentation regarding cooperation, coalition formation and negotiation. Immediate future work will involve the completion of experimentation using existing coordination and coalition schemes.

References

1. Sparck Jones, K., Willett, P.,(Eds.) 1997. *Readings in Information Retrieval.* Morgan Kauffman Publishers Inc. San Francisco, California.
2. Baeza-Yates, R., Ribeiro-Neto, B., 1999. *Modern Information Retrieval.* Addison-Wesley ACM Press, New York.
3. O'Riordan, C., Sorensen, H., 1999. *Information Retrieval and Filtering: An Overview,* in Dept of IT Journal, NUI, Galway.
4. Nwana, H., Ndumu, D., 1995. *An Introduction to Agent Technology,* in Journal of BT Labs.
5. Hewitt, C., 1997. *Viewing Control Structures as Patterns of Passing Messages,* in Artificial Intelligence.
6. Salton, G., Wang, A., Yang, C., 1975. *A Vector Space Model for Information Retrieval,* in Journal of the American Society for Information Science, volume 18, pages 613-620.
7. Huhns, M., Singh, M., (Eds.) 1998. *Readings in Agents.* Morgan Kaufmann Publishers Inc, San Francisco, California.
8. Weiss, G., 1998. *Learning to Coordinate Actions in Multi-Agent Systems,* in Huhns, et al. Readings in Agents.
9. Doran, J., Franklin, S., Jennings, N., Norman, T., 1997. *On Cooperation in multi-agent systems,* in The Knowledge Engineering Review, 12(3).
10. Wooldridge, M., Jennings, N., 1996. *Towards a Theory of Cooperative Problem Solving,* in Proceedings of Modelling Autonomous Agents in a Multi-Agent World (MAAMAW-94).
11. Nwana, H., Lee, L., Jennings, N., 1996.*Coordination in Software Agent systems,* in Journal of BT Labs.
12. Shehory, O., Sycara, K., Jha, S., 1998. *Multi-Agent Coordination through Coalition Formation,* in Proceedings of Agent Theories Agent Languages and Mechanisms.
13. Zlotkin,G., Rosenschein, J., 1994. *Coalition, Cryptography and Stability: Mechanisms for Coalition Formation in Task Oriented Domains,* in Proceedings of the 12th National Conference on Artificial Intelligence.
14. Axelrod, R., 1984. *The Evolution of Cooperation.* Basic Books, New York.
15. Tuomela, R., 2000. *Cooperation: A Philosophical Study.* Philosophical Studies Series, Kluwer Academic Publishers.
16. Wooldridge, M., 1995. *Intelligent Agents: Theory and Practice,* in Knowledge Engineering Review.10(2).
17. Callan, James P., Zhihonh, Lu W., Croft, B., 1995. *Searching Distributed Collections with Inference Networks,* in Proceedings of the 18th Int. ACM SIGIR Conference on Research and Development in Information Retrieval
18. Voorhees, E., Siemens Trec-4 Report. *Further Experiments with Database Merging,* in Harman, D K., Proceedings of the 4th Text Retrieval Conference, 1998. pages 121-130.
19. Sparck Jones, K., Walker, S., Robertson, S., 1998. *A probabilistic model of information retrieval: development and comparative experiments,* in Information Processing and Management, 36(6).