Finding Partners to Form Information Sharing Networks in Open Multi-Agent Systems

K. Suzanne Barber, Jisun Park
The Laboratory for Intelligent Processes and Systems
The University of Texas at Austin
201 E. 24th St. ACES 5.402
Austin, TX, 78712, USA
{barber,jisun}@lips.utexas.edu

Abstract

Information quality assurance despite the existence of uncertainty about sources can be investigated in the context of soft security where an agent maintains trustworthiness evaluations of its information sources to assist in the evaluation of incoming information quality. Dependency inherently exists in a system where agents do not have self-sufficient sensing or data collection capabilities. In other words, an agent requires information from others to achieve its goals. Finding the best partners to form an information sharing network using the notion of trustworthiness can be triggered by dependency adaptation. The research experimentally demonstrates that dependency adaptation can lead a system to form an information sharing network which results in high efficiency and information quality.

1. Introduction

Information quality assurance under the existence of uncertainty can be investigated in the context of soft security (Rasmusson and Janson 1996), where an agent maintains trustworthiness evaluations of its information sources to assist in the evaluation of incoming information quality from those sources. Various trustworthiness evaluation mechanisms have been proposed to handle the uncertainty of both information and information sources (Dragoni and Giorgini 1997; Schillo, Funk et al. 2000; Barber and Kim 2003; Barber and Park 2003; Falcone, Pezzulo et al. 2003). Trustworthiness evaluation can be used to find the best partners from whom to gather information. In this paper, the relationship between dependency and trustworthiness is contemplated to realize best partner selection with respect to efficiency and information quality assurance.

Dependencies exist between agents (and information sources) when each agent cannot be complete in its information acquisition capability and/or in tasks it can fulfill to achieve goals. If an agent is dependent on other agents, it is dependent on them with respect not only to capability, but also to the reliability. Depending on an unreliable entity can cause an agent to become unreliable. Therefore, it is necessary to find reliable partners to depend on.

In addition, if the dependencies are reciprocal, the agents can form a coalition so that they can work together by exploiting each other efficiently. Even when the dependencies are unilateral, the depending agent can form a subjective coalition by filtering out bad partners, meaning that only the dependent agent is concerned about the coalition and the other members do not care if the dependent agent considers them as members of the coalition or not. When forming a coalition agents should be able to distinguish between good and bad partners in the sense that good partners provide required information and service in a reliable way. The resulting coalition should be able to change to adapt dependencies to cope with the dynamics of information source trustworthiness.

According to (Schillo, Burckert et al. 2001), robustness is the ability to maintain "safety-responsibilities" (Wooldridge, Jennings et al. 1999) even with the occurrence of disturbing events. In other words, robustness needs to be related to faults in systems. In MAS, openness and uncertainty necessitate different kinds of faults maintenance. Various replication schemes (Goodmand, Skeen et al. 1983; Davcev and Buckhard 1985; Budhiraja, Marzullo et al. 1993; Schneider 1993) have offered significant advances for fault-tolerance in traditional distributed systems, but they do not work in the face of maliciousness and innocuous quality degradation which are typical in open systems. In contrast with traditional fault-tolerance approaches, robustness can be enhanced by forming a coalition with reliable partners, not in the way that agents confront the faults but by reducing the possibility of fault occurrence.

Goals impose resource (information and service) requirements on each agent. This means that a set of resources are required for an agent to achieve its goals. Therefore, satisfying information and service requirements is a necessary condition for goal achievement. When agents form an organization (or team, coalition, group, etc) to achieve some goals, each agent should be able to get the required information and service from the organization to achieve the intended goals. An ideal organization satisfies the resource requirement of member agents and has a maximum robustness in a given situation.

In information systems where the required resource is information, reliability is represented by trustworthiness of the provided information and/or information sources. Untrustworthy information can confuse or degrade the quality of outcomes from information processing agents.

The goal of this research is to enhance reliability and robustness of an open MAS by deploying the notion of trustworthiness as well as resource dependency, thus to assure the quality of information despite the level of uncertainty surrounding both information and sources. In particular, two main issues addressed by allowing an agent to adapt its dependencies are: *1) search space reduction* and *2) best partner selection.*

The paper is organized as follows. In the next section, the problem of interest is described. Section 3 and 4 explain how dependency adaptation and trustworthiness can be related for partner selection. Section 5 shows experimental results performed in dynamic environments. Section 6 concludes the paper.

2. Problem Description

When an agent requires information from external sources, the agent is dependent on both the information and the sources. $\Phi_a(s)$ is the abstract representation of an information provider *s*'s reliability (trustworthiness in this paper) from agent ***a***'s perspective. $R_a = \{r_1, r_2, ..., r_n\}$ is the set of information agent ***a*** requires. The *Information Pool (IP)* of agent ***a*** is a set of all tuples $< r_k, s, \Phi_a(s) >$, where *s* is a provider of r_k. The *Information Combination Pool (ICP)* of agent ***a*** is a set of tuple sets, where each tuple set $X_j = \{< r_k, s_j, \Phi_a(s_j) > \mid r_k \text{ is provided by } s_j\}$ is a set of tuple combinations which satisfies the information requirements ($\bigcup_{\forall k, < r_k, s_j, \Phi_a(s_j) > \in X_j} r_m = R_a$).

When an agent needs any information, the agent constructs a relationship (i.e. dependency) with the sources of the information. Unilateral relationships are specified by one end of the provider-consumer pair. If a consumer filters out bad providers and the providers are not concerned about it, the relationship is unilateral. In this case, what the consumer evaluates about the providers dominates the relationship determination. Mutual relationships can be constituted by agreement among the stakeholders (i.e. both providers and consumers). For most cases, mutual relationship can be realized when the existence of the dependency increases the benefits for involved participants. In a system where the primary resource is information so the resource requirements are limited to information, information filtering is typically a unilateral relationship. An information consuming agent may select information and information sources based on an evaluation of the providers, but the provider does not care if the consumer takes the information it provides.

2.1. Information Filtering and Information Sharing Network

Information filtering in an Information Sharing Networks (ISN) can be achieved by unilateral dependency adaptation, where information requirements are satisfied and the quality of accepted information is maximized. In the case where one source can provide multiple elements in the set R_a that are required by agent ***a*** and there also exist multiple sources for multiple elements in R_a, dependencies can be distributed among as many sources as possible or can be concentrated on a minimum number of sources, assuming the trustworthiness of respective sources are not much different. The former is diversification of sources and the latter is centralization of information sources. (e.g., agent ***a*** requires information r_1, r_2, information source 1 provides both information r_1 and r_2, information source 2 provides information r_1, information source 3 provides information r_2, the possible combinations of information sources include ({1},{1}), ({1},{3}), ({2},{3}), ({1,2},{3}), ({1,2},{1,3}), etc, where (X, Y) means agent ***a*** receives information r_1 from X, and information r_2 from Y). There exist tradeoffs between diversification and centralization of information sources. Providing dependency to be a quantity increasing as the number of information and/or the number of sources increase, diversification increases the dependency. It means that ISN is more likely to be effected by the dynamics of the system but such effect is relatively small. Centralization decreases the degree of dependence with respect to the number of sources and means that ISN is less likely to be affected by the system dynamics but such effect is relatively large.

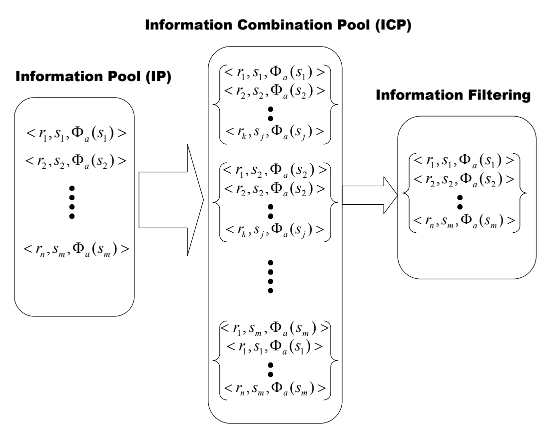

Figure 1. Process of information sharing network formation by information filtering

Figure 1 depicts the process of information filtering. There are two main concerns in the process. The first is to build the *Information Combination Pool* (ICP) from the *Information Pool* (IP). The size of ICP increases exponentially as the amount of information ($|R|$) and corresponding sources ($|S|$) increase. The second is to find the best (or near-best) set of information and sources from ICP. As mentioned above, trustworthiness of sources and degree of dependence distribution are considered as filtering criteria. This paper assumes we have a trustworthiness evaluation mechanism (Barber and Kim 2003) and an agent already knows the potential information sources. The issues addressed in this paper can be summarized as 1) *How to reduce the size of Information Combination Pool (ICP)* and 2) *How to find the best set of information sources given* R *and IP*.

3. Information Combination Pool (ICP) and Information Dependency

The size of ICP depends on the agent's information requirements as well as the number of sources satisfying the information requirements. Let ***N*** be the number of information sources ($|S|$), $n(r_i)$ be the number of potential sources for information r_i. If an agent requires ***M*** information ($M = |R_a|$), the possible number of information and information source combinations (|ICP|) is $\prod_{i=1}^{M}\left(\sum_{j=1}^{n(r_i)} {}_{n(i)}C_j\right)$, which increases exponentially as the number of information required increases. As an example, suppose agent ***x*** requires information $\{r_1, r_2, r_3\}$, information source 1 provides $\{r_1, r_2\}$, source 2 provides $\{r_2, r_3\}$, and source 3 provides $\{r_3, r_1\}$ (Figure 2). In this case, total number of possible combinations is 27. Comparing all the elements in ICP is a simple way to find the best ISN, but it is significantly complex in computation and memory.

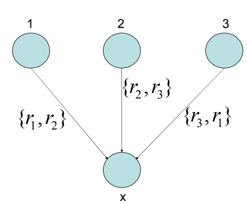

Figure 2. Potential information sources

The approach proposed in this paper is to relate resource dependency to the ISN. It is reasonable to assume that information dependency is inherent since all the agents can not have perfect information acquisition capabilities. Information dependency can be quantified by the following equation, where $N_p = |S_p|$, S_p is a set of sources which are selected as information sources ($N_p \le N$), also $n_p(j)$ is the number of selected information sources for information r_j

$$Dependency(a) = \frac{N_p \sum_{all\ j, \text{s.t. } r_j \in R_a} n_p(j)}{\sum_{all\ k, \text{s.t. } r_k \in R_a} n(k)}$$

For example, the dependency from Figure 2, assuming agent ***x*** filters out information from source ***b***, is 4 x 2/6 = 1.33. The implication of the dependency metric is that it can handle open environments and dependency is represented in terms of information itself and the information sources. Therefore, dependency increases if an agent receives more information from a fixed number of sources as well as if an agent receives information from more sources. Since dependency is related to the number of information an agent receives and the number of information sources for that information, dependency can be restricted to a certain range to reduce the size of the ICP space. Figure 3 shows an unreduced ICP which is represented by a lattice derived based on the dependency definition and built by adding/removing only one source for each edge. Each node except for top and bottom nodes has both parents and children; Parents have higher dependency values and children have lower dependency values. The bottom node is not reachable because it does not satisfy the agent's information requirements. Parents of a given node represent a minimal change in source combinations while increasing the level of dependence. Children of a given node represent a minimal change in source combination while decreasing the level of dependence. Therefore, parents and children are a set of information-source combinations which can be reached by minimum combination change.

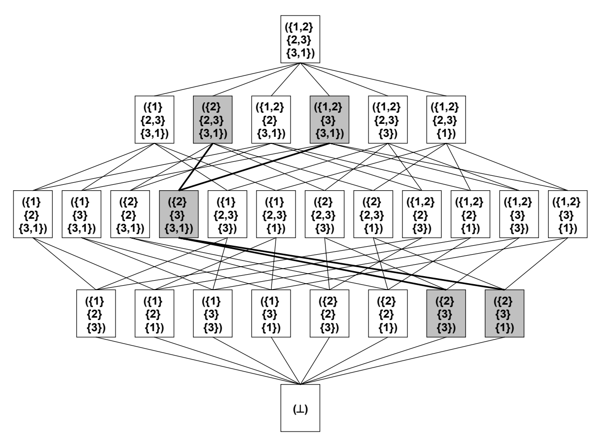

Figure 3. Information Combination Pool built from Figure 2

4. Finding Partners by Adaptation

At a given state, if an agent has to decide whether there exist better combinations of information sources on which to rely there can be different ways to determine the existence of those combinations. Trustworthiness accumulation is used as a metric for triggering dependency adaptation. If the sum of trustworthiness decreases for a current ISN, current information sources are not the best combination, so an agent needs to find a better combination by changing dependencies. The direction of dependency change (i.e. increasing or decreasing the number of information or the number of information sources) plays an important role for reaching a better or the best combination. Additionally, if the system is not stable, meaning the trustworthiness of information sources fluctuates over time, dependency adaptation can direct an agent to a better combination of information sources.

Threshold based dependency adaptation is to decide the direction of dependency change by establishing a threshold. If the current dependency is above the threshold and the dependency needs to be changed, decrease the dependency toward the threshold. Obviously, there are problems with this approach, which include the potential oscillation along the threshold and difficulty in deciding the threshold a priori. The other approach taken in this paper is to use history-based credit assignment. An agent starts with no filtering – meaning maximum dependency. If the sum of trustworthiness decreases, the agent starts to decrease its dependencies since that is the only direction available at that point. Afterwards the direction of dependency change is determined by comparing the values of the Increase Utility and the Decrease Utility. The Increase Utility value is incremented by the predefined incremental credit value (**α** or **δ)** when increasing dependency produces an increase in the sum of trustworthiness or decreasing dependency produces decrease in the sum of trustworthiness. The Decrease Utility value is increased by the predefined incremental credit value (**β** or **γ)** when decreasing dependency produces increase in the sum of trustworthiness or increasing dependency produces decrease in the sum of trustworthiness. Figure 4 summarizes the rules for crediting utilities toward both directions. At a given time, if the previous direction makes the sum of trustworthiness increase, increase the utility for the current direction so that it keeps moving toward the same direction, and vice versa.

The Dependency direction utilities decide to which direction to search in the ICP to find a better combination of sources. Given a node, the lattice can be expanded from the node toward each direction without traversing all the nodes in the search space. For example, consider an agent whose current sources are ({2} {3} {3,1}), i.e. the shaded node in the middle row of Figure 3. If the sum of trustworthiness of those sources decreases over time, then the agent decides to change dependency. If the Increase Utility has a larger value than the Decrease Utility value up to that point, the agent should move in the direction of increases dependency and choose one of its parents – ({2} {2,3} {3,1}) and ({1,2} {3} {3,1}). The best combination is selected from the reduced search space by comparing the sum of trustworthiness of the sources. The speed of adaptation and the path to the best combination of information sources depends on the utility incremental values (**α**, **β**, **γ**, **δ**).

```
If (DependencyDirection==INCREASE &&
    Sum(Trustworthiness) == INCREASE)
    IncreaseUtility += α;
else if (DependencyDirection == DECREASE &&
    Sum(Trustworthiness) == INCREASE)
    DecreaseUtility += β;
else if (DependencyDirection == INCREASE &&
    Sum(Trustworthiness) == DECREASE)
    DecreaseUtility += γ;
else if (DependencyDirection == DECREASE &&
    Sum(Trustworthiness) == DECREASE)
    IncreaseUtility += δ;
```

Figure 4. Guards for utility functions

5. Experiments

Experiments were performed in a target tracking domain. In this domain, we have moving targets and information sources which track the location of the moving targets. The moving targets are assumed to move forward in 1-D space, so the locations are represented by a single floating-point number. Information sources are assumed to be imperfect meaning there exist errors in the locations provided to the agent. There may exist bad sources which have relatively higher errors in the target location information they provide.

In the experiments, an agent tracks the location of 20 moving targets, and there are 30 information sources with each providing the locations of 15 targets. Mean Square Error (MSE) is used for a performance metric. The locations reported have normal distributions with a mean value equal to the true location and variance proportional to the timestep. Since a location is represented by a single floating-point number, a reasonable base line for not using ISN is to average all the inputs. We have four schemes for ISN adaptation.

- ISN 1: threshold-based dependency adaptation without trustworthiness discredit for non-selected sources
- ISN 2: history-based credit assignment without trustworthiness discredit for non-selected sources
- ISN 3: threshold-based dependency adaptation with trustworthiness discredit for non-selected sources
- ISN 4: history-based credit assignment with trustworthiness discredit for non-selected sources

Figure 5 shows a case where there exist no bad sources. 5th-order polynomial regression trendlines of MSEs are drawn for each scheme. In this case averaging shows the best performance with respect to MSE. However, ISN also shows reasonably acceptable values considering the scale of the MSE. The differences between the MSE for averaging all the inputs and the MSE for any of ISNs differ according to the utility incremental values. For example, if increase utility value gets more credit (large α,δ) the ISN tends to increase the dependency so less sources are excluded. In this experiment all the utility incremental values are set to be equal.

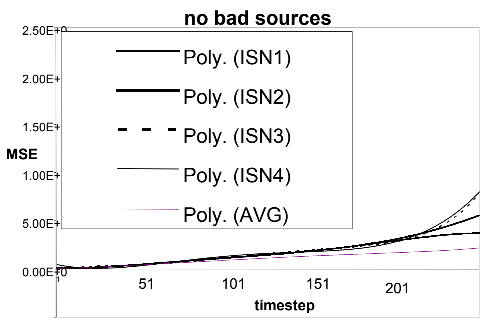

Figure 5. MSE without bad sources

In Figure 6, 20% of sources are bad from the beginning to the end. Bad sources provide location values which have a normal distribution with a mean of ½ of the true location and a variance proportional to the timestep. For the averaging case, the MSE increases as the timestep increases since variance increases. However, ISNs show the MSE increases in the initial stage but the MSEs decrease so that the difference between MSE for averaging and that of ISNs becomes significant. This is because the agent adaptively changes the dependency and excludes bad sources.

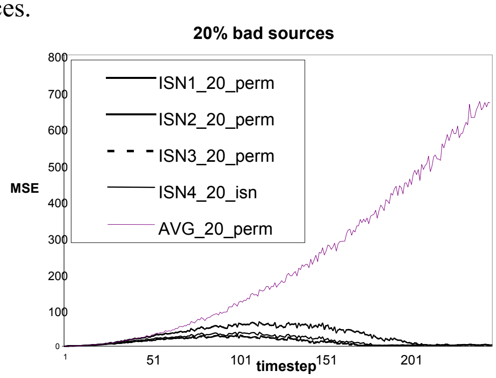

Figure 6. MSE with 20% bad sources

In case of intentionally malicious information sources or innocuously unreliable sources, it is possible for the sources to provide reliable information and unreliable information alternately. Figure 7 shows the case where 20% of information sources are bad sources so they provide unreliable information by reporting ½ of the true location value they sense up to timestep 99. Between timestep 100 and 200, all sources are providing reliable information so we can see the MSE decreases in this time window. At timestep 201, the 20% sources turned into bad sources again so the MSE increases especially for averaging case. Even in this dynamic case, ISNs adaptively decrease MSE by finding the best partners.

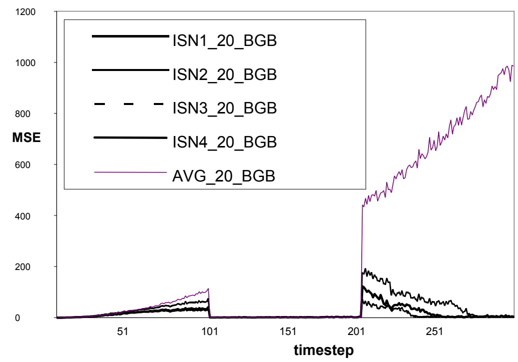

Figure 7. MSE with 20% bad sources with a short time window exception

Figure 8 and 9 show the opposite cases, where bad sources supply ½ of the true location values for timestep 100 to 200. The averaging scheme shows an increase in MSE but ISNs adaptively find better partners to reduce the MSE. The adaptation speed for ISN2 and ISN4 depends on the utility incremental values. In Figure 8, decrease credits (β, γ) are larger than increase credits (α, δ). This means that the agent tends to decrease dependency when dependency adaptation is necessary. This setup is more efficient when there is a higher possibility of the existence of bad sources while it increases the MSE when there are no bad sources as shown in the Figure 8. This is called pessimistic dependency adaptation. In Figure 9, increase credits (α, δ) are larger than decrease credits (β, γ). In this case, adaptation speed is slower than the opposite case but it reduces MSE when there are no bad sources. This is called optimistic because it tends to increase the dependency which means it is more suitable for the no bad source case.

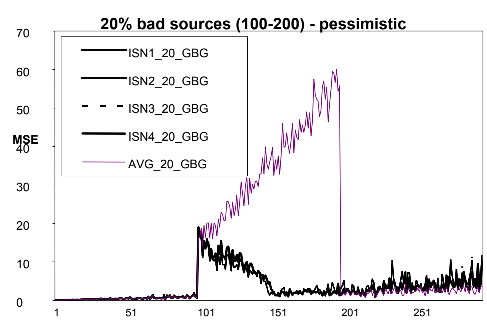

Figure 8. MSE with 20% bad sources in a short time window (pessimistic parameter settings)

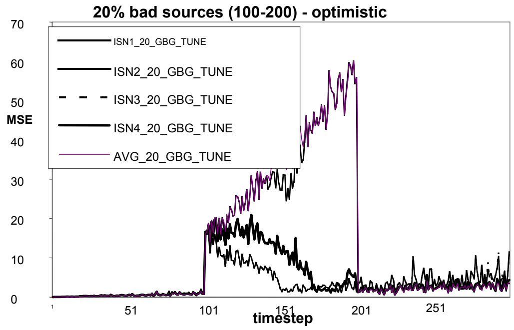

Figure 9. MSE with 20% bad sources in a short time window (optimistic parameter settings)

6. Conclusions

For information quality assurance in open environments, it is paramount that decision-makers (agents) find the best, most reliable providers for the information on which they depend. This research experimentally demonstrates that information and source dependency plays a significant role in forming the best information sharing networks. Dependency adaptation (i.e. selecting a better or the best source combinations) reduces search space so that partner selection can be performed on large-scale systems and enables bad source isolation in an efficient way. Since dependency adaptation continuously looks for a better solution from a small search space by expanding the space effectively, proposed methods for forming information sharing networks show marked improvements over simply selecting a single most trustworthy source.

7. Acknowledgement

This research was supported in part under a subcontract with ScenPro, Inc. funded by U.S. Department of Defense, Small Business Technology Transfer (STTR) Program, Contract Number F30602-99-C-0193.

References

Barber, K. S. and J. Kim 2003. Soft Security: Isolating Unreliable Agents from Society. *Trust, Reputation, and Security: Theories and Practice*. M. Singh, Springer**:** 224-234.

Barber, K. S. and J. Park 2003. Autonomy Affected by Beliefs: Building Information Sharing Networks with Trustworthy Providers. In Proceedings of workshop on Autonomy, Delegation, and Control at the 2nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS2003), Melbourne, Australia.

Budhiraja, N., K. Marzullo, et al. 1993. The Primary-Backup Approach. *Distributed Systems*. S. Mullender, Addison-Wesley**:** 199-216.

Davcev, D. and W. Buckhard 1985. Consistency and Recovery Control for Replicated Files. In Proceedings of ACM Symposium on Operating Systems Principles.

Dragoni, A. F. and P. Giorgini 1997. Learning Agents' Reliability through Bayesian Conditioning: a simulation study. *In Proceedings of Learning in DAI Systems,* Springer-Verlag.

Falcone, R., G. Pezzulo, et al. 2003. A fuzzy approach to a belief-based trust computation. *Lecture Notes in Artificial Intelligence*. 2631**:** 73-86.

Goodmand, N., D. Skeen, et al. 1983. A Recovery Algorithm for a Distributed Database System. In Proceedings of the 2nd ACM SIGATC-SIGMOD Symposium on Principles of Database Systems.

Rasmusson, L. and S. Janson 1996. Simulated social control for secure Internet commerce. *In Proceedings of New Security Paradigms '96*, ACM Press.

Schillo, M., H. J. Burckert, et al. 2001. Towards a Definition of Robusness for Market-Style Open Multi-Agent Systems. In Proceedings of the Fifth International Conference on Autonomous Agents (Agents-01).

Schillo, M., P. Funk, et al. 2000. Using Trust for Detecting Deceitful Agents in Artificial Societies. *Applied Artificial Intelligence Journal, Special Issue on Deception, Fraud and Trust in Agent Societies*: 825-848.

Schneider, F. B. 1993. Replication Management using the State-Machine Approach. *Distributed Systems*. S. Mullender, ACM Press - Addison Wesley**:** 169-197.

Wooldridge, M. J., N. R. Jennings, et al. 1999. A Methodology for Agent-Oriented Analysis and Design. *In Proceedings of the 3rd International Conference on Autonomous Agents (Agents-99)*, Seattle, WA, ACM Press.