CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

The growing rate of complexities of software systems is getting high with respect to the increasing and unpredicted demands, particularly for distributed systems in the modern technological era. These non-deterministic natures of demands thrive the researchers to discover more sophisticated and flexible technologies to aid the developers in various directions. In this progression, agent technologies and its successors have generated lot of excitements in the recent years because of its assurance as a new successful paradigm for conceptualizing, designing and implementing complex software systems (Katia 1998). Multi Agent System (MAS) is one among in the series, which is mature enough in handling the challenges in development of distributed systems (Jennings et al 2001).

The advent of MASs has brought together many disciplines and they have given us a new perception to look at the distributed systems and provided an efficient path to build more robust and intelligent applications (Scott et al 2001). In the present scenario, innovative MAS standards and supporting guidelines enable the construction of multifaceted distributed software systems in more classical and expedient fashion. These promotions aspects project the MAS as fascinating technology in the domain of distributed application environments.
On the other hand, these endorsement aspects also magnify the proficiency needed to accomplish, extend and of course validate the MAS quality goals and performance with respect to the anticipated applications (Katia 1998). Particularly, the process of MAS performance evaluation is more essential to assess the effectiveness of the intended technologies. This necessitates more effective performance evaluation methods for MAS environments both from MAS and intended application perspectives. Though numerous performance evaluation schemes are available, irrespective of individual’s capabilities and meritorious claims, it is generally accepted that the quantification based performance evaluation mechanisms always support and contribute to appreciable extent in this regard (Marc et al 2004). For most of the approaches, the outcomes possessing Babylonic mismatch with the anticipated processes and it is a crucial factor in MASs research field, which leads to create artificial thresholds to convince mainstream software developers on behalf of the limited merits of the individual techniques (Danny et al 2008).

From these perspectives, the work presented in this thesis is aimed at analyzing the abilities of the existing MAS performance quantification schemes at different levels in order to minimize the degree of mismatch between them and thereby to offer an enhanced scheme for the same with proper justifications.

1.2 PERFORMANCE EVALUATION IN MAS
1.2.1 Characteristics of MAS

Multi Agent System (MAS) is the collection of autonomous software entities (agents) that are determinant and able to exhibit complex flexible behavior in order to achieve their predefined and ad-hoc objectives. The achievements are due to the interactions of the individuals with one another
using standard specifications and recommendations in the dynamic and uncertain environments. This flexible behavior is significant in making multiagent technology evident and also makes it hard to obtain confidence that the system will work as desired (Rafael et al. 2006).

Apart from the inherent characteristics of the autonomous agents, coordination, collaboration and cooperation nature of the MASs reflects the decentralized nature of modern distributed systems. Katia (1998) presented the essential characteristics of MASs: each agent has incomplete information or capabilities for solving the problem and, thus, has a limited viewpoint; there is no system of global control; data are decentralized; and computation is asynchronous. Based on these, the MAS environments can be considered as the natural extension to the philosophy of modularization and encapsulation for the systems of different stakeholders (Franco et al. 2003, Tennenhouse 2000, Parunak 1997). Modularity allows the decomposition of a complex problem and coordination between the modules helps to manage possible interdependencies in the complex problem (Jose et al. 2007). These key properties lead to delegating control to autonomous components within the scope of the specified environment.

In addition to these, the appropriate level of dynamic nature of multiagent interactions are very much suitable to the open ended systems in which the member components and their interface outlines are to be constantly changed (Estrin et al. 2002, Ripeani et al. 2002, Ricci et al. 2002). The property of openness of distributed application circumstances results in impossibility in define priori all potential interdependencies between the member components. Autonomous components delegated of their own control can be enhanced with refined social abilities that makes the system to capable enough to make the decisions about the scope and nature of their interactions at run-time in order to support environment to achieve the goal of the overall system.
Decentralization in control leads clear distinction between the intended autonomous entities of the overall system and the local & global resources may offer a sophisticated modeling of the complex problem domain.

In summary, the key features of MAS that tend to attract and motivate the researchers and practitioners towards this exciting field of MAS can be briefly stated as in (Katia 1998) as follows:

- MASs are capable provide effective solutions for more complex problems, which are more critical for single agent or centralized systems.
- MASs allow the interconnections and interoperations of multiple existing legacy systems and this property avoid the process of completely rewriting such software, which is prohibitively expensive and is often simply impossible.
- MASs are to afford solutions to the problems that can naturally be regarded as a society of autonomous interacting agents that efficiently use information sources that are spatially distributed.
- MASs can offer the solutions to the situations where the distributed expertise is inevitable.

In spite of these meritorious qualities, MASs requires enhanced expertise in design, development, validation and maintenance phases in a comprehensive fashion. These requirements motivate the researchers strive to come across with able techniques to manage MASs throughout its life cycle. In this view, this thesis presented an enhanced scheme of MAS performance evaluation based on unique set of measurement schemes.
1.2.2 Different Schemes of Performance Evaluation in MAS

The idea of performance evaluation in MAS is not a new one, several research proposals and recommendations have already been made on verification and validation of MASs (David 2001, Mark et al 1999, Tony et al 2002, Poutakidis et al 2002, Steven et al 2002). An important challenge or crisis in the field of MAS performance evaluation is the development of quantification methods to assess the significant properties or attributes of MASs. Such methods could be used to present an in-depth analysis of the MASs providing feedback prior to the deployment.

Furthermore, measurement based assessment models could offer well-defined guidelines for the design of MASs that will ensure that the system will conform to the required properties. Based on the environmental qualities and necessities, in general, MAS performance evaluation schemes are categorized into three different types (Marc et al 2004); theorem proving, model checking and testing.

Theorem proving approach: Theorem proving techniques are originated from logic systems and concerned to verify the set of procedures that are to satisfy the goals and objectives of the estimated system through deductions and inferences (Marc et al 2004). However, the process of finding the proofs may be difficult to perform especially when the application domain is extremely distributed with large volume of interactions between agents (Denis et al 2006).

Model checking approach: Model checking techniques are intended to derive a model of the projected system and verify the temporal logic formulae holds for the model (Clarke et al 2000, Marc et al 2004). Model checking costs time and the problem must be reduced to a finite one, with a limited number of
states, which is not possible in most of the cases, particularly MAS environments (Denis et al 2006). Rafael et al (2006) explored the significant aspects of model checking in a very clear and descriptive fashion: Generally, technologists prefer model checking compared to theorem proving because it is easier to use and more importantly, it can provide counter examples when the system fails to gratify a preferred behavior; Most of the work done on model checking in the view of evaluating the performance of MASs is quite theoretical in order to verify the properties of specific aspects of a MAS; The typical nature of MASs necessitates the model checking techniques that are to be tailored particularly for MAS; On the other hand, the indecisive and uncertain characteristics of MAS specifically can be made to expect more efficient verification methodologies than normally expected in traditional software, still the expectation continues.

*Testing approach:* Testing is the process to get the confirmation over the behavior of the proposed system with respect to the test cases corresponds to the anticipated system (Marc et al 2004). Contrary to the methods of theorem proving and model checking, tests are not required to be so exhaustive. As a consequence, the confidence brought by tests is condensed, but is easier to obtain (Denis et al 2006). This category includes debugging in MASs also. For any software system, it is indispensable that when the system is installed and used, there should be assurance that it must do what it is hypothetically to do. Typically, this confidence is achieved through testing.

However, for MASs that are able to exhibit flexible behavior, achieving their goals in a range of ways depending on the situation, it is harder to achieve confidence in the system through testing and hence there has been a rather limited amount of work on testing agent systems, but on the other side, there has been interest in using formal methods, especially model checking, to verify agent systems (Rafael et al 2006).
In summary, in order to offer significant support for guaranteeing that MAS performs perfectly, much work is needed in theorem proving, model checking and testing approaches and tools (Rafael et al 2006). More grippingly, approaches that combine these three approaches are also likely to materialize for MAS, as they have eligibility to claim for better consolidated performance in the anticipated environments. For example, when full fledged verification is not possible because the system’s state space is too large, even after the use of state-space reduction techniques, practical model checking tools can still be used, for example, to help find the required input leading to special cases that can be potentially useful in testing a system (Rafael et al 2006).

It is found that metrics based evaluations can serve as a platform to combine all these three schemes or otherwise they can serve as the assistive technologies for all these three schemes (Marc et al 2004, Dariusz et al 2008). In this perspective, the work narrated in this thesis is focused on evaluating MAS performance using a new set of metrics, which might be a supportive proposal for the last two evaluation schemes; model checking and testing based MAS performance evaluations.

1.2.3 Need for Metrics Based Performance Evaluation in MAS

Generally, it is also widely accepted that the measure of plan quality is one of the important characteristics of MAS simulators and platforms (Martha et al 1993). It is also stated that, measures of plan quality are used by agents to determine if they are going to achieve their goal, but should not be of direct concern to the simulator. Hence, it can be claimed that the performance evaluation schemes of MAS should be primarily oriented towards the application and little bit concern may be offered to the methodologies and frameworks, but not towards the platforms and tools.
On the other side, it can also be claimed that proper measurement theories must incorporated with any of the performance evaluation schemes rather than the qualitative descriptions. Irrespective of the domains, metrics based evaluations are always significant, because if they were confirmed by experiments, it would allow them to be used to valuate many other realizations without actual tests implementations (Dariusz et al 2008). The primary advantage of this approach is to combine any of the core performance evaluation schemes of MAS. This is also applicable for evaluating both deterministic and non-deterministic attributes of MAS by quantifying them with appropriate amendments.

The notable significant feature associated with metrics based evaluation is two levels of verification; first to verify a correlation between measurements done during tests and second confirm the standard of the performance estimations exclusively based on projected metrics. As discussed, metrics based evaluations have the ability of serving as the assistive technologies for all the three schemes of MAS performance evaluations; theorem proving, model checking and testing (Marc et al 2004, Dariusz et al 2008). These meritorious aspects strongly support the need for metrics based MAS performance evaluations.

1.3 MOTIVATIONS

The motivation of the work reported in this thesis has two folds; from the MAS performance evaluation point of view and the software measurement theory point of view. In general, MAS performance evaluation frameworks may be built as either application domain dependent or independent. In case of domain independent approaches, the evaluation schemes may be served as blueprint and lead to commit appropriate adjustments in consideration to the respective applications. In spite of this limitation, these models are capable
enough to serve as decision support tools to select better mechanisms for the pre-planned processes only. On the other hand, domain dependent schemes may be served as evaluation models with implicit proof making platforms (i.e. readymade space dimension to prove the claims of the proposals).

Apart from these discussions, irrespective of the model-specific features, it is also stated that, if the dynamics of the MASs are properly validated in all required environment states, then it is possible to utilize and establish the merits of MASs at the required levels appropriately. This claim can be made only based on the available support of level specific validation methodologies of the MASs. But the availability of level or attribute specific validation methodologies and guidelines are so limited (Ingo e al 2009) and this potential insufficient status motivated the researchers to find better validation models for MASs. It is also discussed that, irrespective of the evaluation models and domains, the measurement process is apparent in the MASs and almost all the research works are centered on this either to contribute to it or to make use of it. The scope for extending the measurement strategies in MASs is highly evident and this should be applied throughout the life cycle of a MAS. But majority of the works suffered with Babylonic mismatch between the proposed evaluation schemes and the anticipated objectives.

From these it can be observed that the need for refined and fine grained measurement strategies in MASs is very important and it can be extended to various possible levels such as independent and dependent load execution strategies, negotiation, interaction and communication schemes, etc… So it is realized that a framework of fine grained measurement schemes for MASs is necessary at this juncture. Hence, from measurement based MAS performance evaluation point of view, the work reported in this thesis is motivated by the fact that a fine grained measurement framework seems to be
a better approach for quantifying and evaluating the MAS performance attributes.

On the other perspective, in general, appropriate measurement theory and metrics are highly significant in software systems. Irrespective of the models and domains, the measurement process is an inevitable component in the software science and almost all the research works are centered on this either to contribute to it or to make use of it. Norman et al (2001) claimed the significances of measurement theory particularly in software systems and those claims can be expanded as follows:

- the objective of measurement is to quantify an attribute of the software system in order to have the intentional control and management over the same.
- Fine grained measurement strategies capture and quantify the precise information about the attributes of the software systems to improve the interpretations over the influence of the same.
- the accuracy of a measure is depends upon two important key factors; quality of the measuring instrument as well as the definition of the measurement with respect to the attributes to be measured.
- indeed, measuring the immeasurable aspects will improve our understanding of particular attributes of the software systems.
- even when it is not clear about measurement of an attribute, the act of proposing such measures will open a debate that leads to a greater understanding of the relevant domain, particularly in case of software systems.
From Norman et al (2001) points of views, it is clearly understood that the degree of maturity of interpretations over the impact of individual attributes is directly proportional to the level of granularization of measuring the attributes. It is also expressed that the accuracy of a measure is depends on two key factors; measuring instrument and definition of the measurement. On the other hand, it can also be extended as the accuracy of a measure is also depends upon the ability of making the things to be measured as more measurable and this may be resulted in proper control and management over anticipated attributes.

Hence, from the software measurement theory point of view, these claims motivated this research to proffer software metrics based evaluation framework, which will facilitate the designers and developers to assess the MASs at all required stages. In addition to this, it is also aimed that by proposing a set of new evaluation metrics for MAS, invite the debates and thereby to define the new directions of research in the specified domain in future.

1.4 OUR CONTRIBUTIONS

As discussed earlier, the endorsement aspects of MAS also magnify the proficiency needed to accomplish, extend and indeed to validate the MAS quality goals and performance with respect to the anticipated applications (Katia 1998). Particularly, the process of MAS performance evaluation is more essential to assess the effectiveness of the intended technologies. This necessitates more effective performance evaluation methods for MAS environments both from MAS and projected application perspectives.

Though numerous performance evaluation schemes are available, irrespective of individual’s capabilities and meritorious claims, metrics based
evaluations can serve as the assistive technologies along with any of the core performance evaluation schemes of MAS (Marc et al 2004, Dariusz et al 2008). This is also applicable for evaluating both deterministic and non-deterministic attributes of MAS by quantifying them with appropriate amendments. Hence, it can be claimed that the performance evaluation schemes of MAS should be primarily oriented towards the application and little bit concern may be offered to the methodologies and frameworks, but not towards the platforms and tools.

On the other side, it can also be claimed that proper measurement theories must incorporated with any of the performance evaluation schemes rather than the qualitative descriptions in order to improve the overall valuation procedures. This work has been primarily motivated by the fact that a fine grained measurement framework seems to be a better approach for quantifying the MAS performance attributes.

For achieving the above, this experimental research is being organized into several phases and the major contributions are listed as follows:

- An extensive survey has been made over the recent related works and it has been concluded with the necessity for having an enhanced measurement schemes for quantifying the estimated performance attributes of the MAS environments.
- Five layer experimental framework is being developed to experimentally prove the asserts made in this research work.
- New sets of MAS performance attributes of two different categories are devised along with the appropriate measurement scheme.
• A enhanced multi-agent framework for software testing with automated resource negotiation and allocation schemes has been built as a test bed for the proposed research.
• Designed a set of appropriate experiments such that to prove the claims made in the proposed research in the versatile and realistic environments.
• Applied the statistical techniques to analyze the results obtained from experiments in order to validate the outcomes of the research presented in this thesis.

1.5 ORGANIZATION OF THE THESIS

The remainder of this thesis is organized as follows:
• Chapter 2 presents a concise review on recent contributions related to the work presented in this thesis. This review is organized into two sections based on the theme of MAS performance evaluations. Discussion in first section is focused on metrics based MAS framework performance evaluations, whereas the next section is focused on metrics based MAS application performance evaluations. This chapter also clearly describes the babylonic mismatch between the proposed metrics based MAS evaluation schemes and the anticipated objectives.
• Chapter 3 defines the goals of this research and explains the line of research to achieve the goals defined. This chapter also defines the layered framework of experimentation methodology followed in this research.
• Chapter 4 elucidates the method of deriving communication oriented metrics in different dimensions. The corresponding theorems in this chapter define the associated design and
operation oriented conditions related to the communicational issues of MASs.

- Chapter 5 describes about the coordination oriented metrics in different dimensions and their variants. The appropriateness and necessity of the proposed metrics are also illustrated in this chapter.

- Chapter 6 outlines the test bed constructed for this research in order to validate the proposed metrics models. This chapter also describes the necessity of the having variety of agents at different levels with different abilities.

- Chapter 7 offers quantitative analyses on the experimental results using appropriate analyzing methods including statistical models with corresponding interpretations.

- Chapter 8 provides the concluding remarks of the work presented in this thesis and the future enhancements of the proposed line of research.