Agents: Towards a Society in Which Humans and Computers Cohabitate*  
(Extended Abstract)

Mario Tokoro**
Faculty of Science and Technology, Keio University.  
3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223 Japan

Abstract. Multi-Agent Systems research is a new field. Therefore, no consensus exists on its definition or its purpose, nor does there exist any generally-agreed way of doing research in this field. In this paper I first look back at the streams of Multi-Agent Systems research, comparing them to Distributed Problem Solving, Distributed Artificial Intelligence, Parallel Artificial Intelligence, Distributed Algorithms, Open Systems, and so forth, and I try to show what Multi-Agent Systems really are. Based on this analysis, I then suggest some aims and future directions for research in this field.

1 Introduction

Multi-Agent Systems (MAS) are, as the name suggests, systems that comprise multiple agents. The word “Agent” from the Latin “agāns” to act, is defined as the producer of an effect, an active substance, a person or thing that performs an action, a representative, and so forth. Among these meanings, a representative, or a thing that performs an action, seems to describe best the word as used in MAS research. That is, it means “an individual that performs an action”, so that in its broadest sense, a MAS is “a system composed of multiple individuals which perform actions”.

The essential motivation of MAS research was to answer the question: “Can a higher level task be achieved by cooperation between multiple subsystems, each of which has a lower ability?” This motivation originated principally in two research fields. The first is called Distributed Problem Solving (DPS), a field that investigates the coordination of multiple subsystems as the way to obtain a higher ability. When it is conducted as research related to intelligence, it is called Distributed Artificial Intelligence (DAI). The second is called Agent-Oriented Programming, a field that investigates the methodology used to construct software, especially for Open Systems.

Two related notions, that of a Network Agent — an entity that moves around in the Internet, and the notion of an Interface Agent — an entity that handles...
the interface between human and computer, are important in connection with the above-mentioned two fields because they provide examples of agents that conform to, or use the ideas originating in those fields.

In the following sections, I will review these four research streams, and I will try to suggest some aims and future directions for research in MAS.

2 Distributed Problem Solving and Distributed Artificial Intelligence

Distributed Problem Solving originated with the naive question: “How should a task be divided into subtasks so that they can be processed concurrently in an efficient way?” It is a field that is application oriented and attempts to deal with real practical problems [9][6]. We also call this area Distributed Artificial Intelligence (DAI), because many of its research results have been presented at conferences related to Artificial Intelligence. It is hard to classify the papers from their contents.

In the case where necessary information about other subsystems can be obtained with negligible delay, the field is called Parallel Problem Solving (or Parallel Artificial Intelligence) and corresponds to the research on Concurrent/Parallel Algorithms. In the case where there is a time delay in retrieving information from other subsystems, or in the case where delays on subsystems and communication paths must be considered, they are genuinely in the field of “Distributed” Problem Solving (or “Distributed” Artificial Intelligence). Those studies that assume that subsystems have only incomplete (partial or containing errors) knowledge in the problem domain can be categorized within the above, and are close to the field of Distributed Algorithms.

In DPS or DAI, as each subsystem begins to have the notion of “self,” they begin to exhibit different characteristics: subsystems begin to pursue different goals for their own benefit instead of the common goal. This leads DPS and DAI to Multi-Agent Systems [3]. A key distinction in such systems is that an agent which gives information to other agents may not always necessarily benefit itself. That is, in Multi-Agent Systems, agents always have to maximize their own benefit according to incomplete information about other agents. This behavior of agents is called “rational.”

Generally, in MAS research, there exists the notion that agents pursue the benefits of the entire system, i.e. altruism. However, my position is that the agent should primarily concern itself with its own benefit. It is not however, always the case that ‘short-termism’ is a prerequisite for altruism. Interestingly, ignoring its own short-term benefit may to some extent benefit the entire system which in turn may lead to its own long-term gain and stability. This is the essence of cooperation.

A fundamental proposition of MAS research can be expressed as follows. “Can each agent obtain higher benefit by cooperating with multiple agents mutually than when it doesn’t cooperate?” Here, the framework and strategy for negotiation become the objects of research. The above-mentioned proposition