A Case Study on the Role of Information for Implicit Coordination

Franziska Klügl
Dep. of Artificial Intelligence
University of Würzburg
Am Hubland
97074 Würzburg, Germany
kluegl@informatik.uni-wuerzburg.de

Ana L. C. Bazzan
Instituto de Informática
UFRGS
Caixa Postal 15064
91.501-970 Porto Alegre, RS, Brazil
bazzan@inf.ufrgs.br

Abstract

One of the major research directions in multi-agent systems is the design of coordination mechanisms. The present approach aims to study the effect of providing agents with two types of information to influence the decision-making process: correct and intentionally manipulated. The results show that, although manipulated information leads to the optimal state, the overall result is best when a certain share of agents ignores the given information.

1. Introduction and Related Work

In multi-agent systems the idea of implicit coordination is not new. Social laws for example aim at producing a implicitly coordinated system. However, finding appropriate social laws is a difficult task. Here, we address the way that a control component could influence the behavior of autonomous agents so that a coordinated situation emerges. This influence is based on information that is given to agents with a bias aiming at maximizing the overall utility of the system. We show this in an abstract scenario, in which an agent faces repeated actions selection such as in [1].

This idea of coordination games was also explored by us in [2] under the focus of reinforcement learning with different learning parameters. We now return to these issues to evaluate the effect of information and information manipulation in this kind of binary choice scenario. These issues are important in mechanism design in the sense that there is a tradeoff between highly autonomous agents versus efficency of the whole system.

AAMAS'04, July 19-23, 2004, New York, New York, USA.

2. Abstract Scenario

In the abstract scenario tackled here N agents repeatedly have to decide between two alternative actions (henceforth alternatives for short). At the end of the round, every agent gets a reward that is computed based on the number of agents who selected the same alternative. The agents do not know the reward function or the rewards of other agents; they just know the reward they get. However, their decisions do influence the reward each receives. Rewards for each agent i are computed so that one alternative, namely M (main) is preferred over the secondary one (S). With $N = 900$, the reward function is such that an equilibrium for the distribution of reward occurs when 600 agents select M and 300 select S. In this case, the reward is 500, no matter a agent selects M or S.

A simple model for adaptive choice is detailed in [2]. Agents decide which alternative to select based on the rewards. This calculation is summarized in a "choice heuristic". Practically, it is the probability according to which an agent selects the main alternative (M). In the adaptive scenario the agent updates this heuristic with a given probability according to the rewards he has obtained selecting that alternative. The update of the heuristic considers the rate of reward for a given alternative and is computed dividing the accumulated reward for alternative A_j by the total reward so far. With learning probability of 0.2 (that means, on average, in every fifth round the agents are adapting their heuristic), the agents learn the optimal heuristic of $\frac{2}{3}$ (on average).

2.1. Scenario with Information

We extended that basic scenario so that agents receive a forecast and the decision-making process was performed in two phases. First, agents make their initial selection (first decision) and inform the information provider about their decisions. Based on this information collected from agents, the provider computes the reward for every agent and sends this information back: the agents receive a forecast information regarding the potential reward they would get if they keep their first decision. Then, they have a second chance to actually take their first choice or change it (second decision). Finally the selections are made yielding the actual rewards.

From previous studies [2] we know that providing information to agents that actually use it results in a quite stable situation. However, the outcome is not the optimal one, as too many agents select the main alternative in the second choice.

2.2. Scenario with Manipulated Information

The main difference to the previous scenario (with forecast) is that, here, the information given to the agents is manipulated and biased towards the equilibrium of the systems. For example: if the current distribution is 800 in M and 100 in S, it is clear that the system is out of the equilibrium and therefore, the rewards to the agents are sub-optimal. When the information provider perceives such an out-of-equilibrium situation, it tries to induce agents to come closer to the equilibrium by given manipulated forecast to them. In the particular case above, since there is an excess of 200 agents selecting M, 200 agents must receive a forecast to change to alternative S.

The most interesting situation is when everyone is informed. In this case correct forecast causes a high inefficient distribution of agents [2]. We compare manipulated information with non-manipulated in two cases: without and with adaptive agents. In the former, agents do not adapt (reinforcement mechanism explained above). Figure 1 depicts the number of agents who selected M.

Comparing the curves for non-learning agents, it is easy to see that the fact that agents receive manipulated information does not improve its performance. With correct information too many agents select M; with manipulated information too few do. Both are far from optimal. This is so because the first and the second choices are random.

An improvement in performance can be observed when agents do learn, under manipulated information. Similar results hold for different shares of uninformed agents, because this share is integrated in the way the provider computes how many agents have to receive a manipulated forecast information.

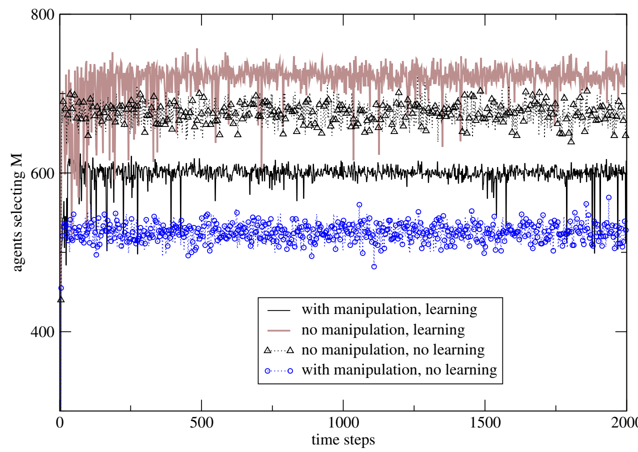

Figure 1. Number of agents selecting the main alternative with and without manipulated forecast information, for learning and non-learning agents.

3. Conclusions and future work

The simulations described here show that, regarding mechanism design, it is useful to manipulate the forecast information given to the agents. Doing so, the information provider is able to divert them to the more convenient alternative, from the point of view of both the overall system and the individual agents (as this is the situation in which individual rewards are the highest). This also holds for a quite large range of assumed share of uninformed agents. Having agents provided with the most accurate information (information about the *actual* state of the system) is not necessarily good. However, the capacity of adaptation (by the agents) seems to be key, together with the capacity of the provider to distribute manipulated information.

The main direction for this research is the design of more complex agent architectures to accommodate competition and self-interest.

References

[1] Arthur, B. Inductive reasoning, bounded rationality and the bar problem. Tech. Rep. 94–03–014, Santa Fe Institute, 1994.

[2] Klügl, F., and Bazzan, A. L. C. Route decision behaviour in a commuting scenario. *Journal of Artificial Societies and Social Simulation 7*, 1 (2004). http://jasss.soc.surrey.ac.uk/7/1/1.html.