COLLABORATION AMONG COMPETING FIRMS
An Application Model about Decision Making Strategy

Marco Remondino and Marco Pironti
e-business L@B, University of Turin, Torino, Italy

Keywords: Collaboration, Bias, Ego biased learning.

Abstract: To understand the adoption of collaborative systems, it is of great importance to know about economical effects of collaboration itself. Decision makers should be able to evaluate potential drawbacks and advantages of collaboration: strategies may be seen as a mixture of cost reduction, product differentiation and improvement of decision making and/or planning. In this context information technology may help a firm to create sustaining competitive advantages over competitors. It is less clear whether collaboration is of any use in such an environment. According to the Economics literature, the most important factors affecting benefits of collaboration are market structure, kind and degree of uncertainty faced by the firms, their risk preferences and the collaboration propensity. The results depend on the way these factors are combined. We present a microeconomic model and use techniques from game theory for the analysis. The way the model is constructed will allow the derivation of closed-form solutions. Traditional learning models can't represent individualities in a social system, or else they represented all of them in the same way – i.e.: as focused and rational agents; they don't represent individual inclinations and preferences. Results indicating whether collaboration in various areas makes sense will be obtained. This makes it possible to judge the potential of available collaborative technology. The basic presented model may be extended in various ways.

1 INTRODUCTION

Decision makers need to understand the economics of collaboration in order to be able to evaluate the potential of collaborative technology. Collaboration among different actors may occur within a firm's boundary or across it. The economic effects of collaboration between firms located along different phases of the value chain (typically supplier-purchaser-relationships) have been extensively studied in the literature.

Usually, transaction cost theory is applied to derive the "optimal" institutional structure (Williamson, 1975). Basic institutional arrangements are hierarchy, market and network cooperation (Clemons et al., 1993). The use of collaborative technology may be especially useful in case of network cooperation. As Clemons et al. point out, the use of IT triggers a move towards such cooperation. As is well-known, strategies of firms may be seen as a mixture of cost reduction, product differentiation and improvement of decision making and/or planning. Information technology may help a firm to create sustaining competitive advantages over competitors. Social models are not all and only about coordination, like iterated games, and agents could have a bias towards a particular behavior, preferring it even if that's not the best of the possible ones. An example from the real world could be the adoption of a technological innovation in a company: even though it can be good for the enterprise to adopt it, the managerial board could be biased and could have a bad attitude towards technology, perceiving a risk which is higher than the real one. Thus, even by looking at the positive figures coming from market studies and so on, they could decide not to adopt it. This is something which is not taken into consideration by traditional learning methods, but that should be considered when modeling social systems, where agents are often supposed to mimic some human behavior. In order to introduce these factors, a formal method is presented in the paper: Ego Biased Learning (EBL).

A first result of the analysis shows that maximization of expected utility may lead to different optimal actions than maximization of expected profits. While the latter in general is a simple optimization problem, maximization of expected utility requires knowledge of the utility

Remondino M. and Pironti M.
COLLABORATION AMONG COMPETING FIRMS - An Application Model about Decision Making Strategy.
DOI: 10.5220/0002796903870391
In *Proceedings of the 6th International Conference on Web Information Systems and Technology (WEBIST 2010)*, page
ISBN: 978-989-674-025-2

function of the decision maker. Note, that this is a more difficult and complex problem.

If costs of information sharing are sufficiently low, information sharing generally is beneficial if development know how is equally distributed. This is an expected result since then development costs may be reduced. This result, however, changes significantly if know how for development is not equally distributed between the competitors. In this case, situations that are well-known from the treatment of “prisoner-dilemma-situations” occur. Results then strongly depend on the degree of risk aversion and the market structure and nearly all “prisoner-dilemma-situations” may be constructed by suitably choosing the model parameters. First mover and follower strategies are then optimal choices depending on risk aversion of the decision makers and market structure.

In some instances, the results obtained will be surprising and contradict rational expectations.

It can be shown, e.g., that the use of information to reduce uncertainty may be harmful for an enterprise. Firms are paid for taking risks. If they try to reduce such risks profits and expected utility may decrease (even if risk is reduced at zero costs, see e.g. Palfrey; 1982).

Once again, such surprising results show the importance of understanding the economic effects of collaboration before deciding on investments in collaboration technologies. But it is very important to consider collaboration expectation as a predominant behavior of other firms.

2 COLLABORATIVE CULTURE

Generally, collaboration among competing firms may occur in many ways. Some examples are joint use of complex technological or marketing processes, bundling products or setting standards. Collaboration typically requires sharing information and know how, as well as resources.

In the literature collaboration problems are usually studied with the help of methods from microeconomics and game theory. It turns out that the most important factors affecting the usefulness of collaboration are the following ones:

- *Market Structure*. If perfect competition prevails collaboration is of limited use. No single firm or proper subsets of firms may influence market prices and/or quantities. In a monopolistic environment there obviously is no room for collaboration. Consequently, the interesting market structure is an oligopoly. Depending on the kind of products offered and the way an equilibrium is obtained, price or quantity setting oligopolies may be distinguished.
- *Product Relationship*. Products offered may be substitutes or complements. In general, we would expect that products of competing firms are substitutes. Product differentiation, however, allows to vary the degree of possible substitution.
- *Distribution of Knowledge and Ability*. The distribution of knowledge and ability is closely related to the possibility of generating sustaining competitive advantages (Choudhury, Sampler; 1997). If a firm has specific knowledge or specific abilities that competitors do not have it may use these skills to outperform competing firms.
- *Kind and Degree of Uncertainty Faced by Competing Firms*. Basically we may distinguish uncertainty with respect to common or private variables. As an example consider demand parameters. They are called common or public variables since they directly affect profits but are not firm specific. On the other hand variable costs are an example of private variables (Jin; 1994, p. 323). They are firm specific. Of course, knowledge of rival’s variable costs may affect a firms own decisions since it may predict rival’s behavior more precisely.
- *Risk Preferences of Competing Firms*. It is assumed that decision makers are risk averse. Hence they will not maximize expected profits as if they were risk neutral but expected utility of profits.

The results obtained depend on the assumptions made about the factors identified above. They partially differ or even contradict each other.

The analysis presented in this paper is carried on with the help of a microeconomic model based on game theory. The basic assumption is that collaboration occurs through knowledge and information sharing, common information collection and/or interpretation. In order to share information, knowledge and know-how, collaborative technology is usually applied. Joint application development and joint use of resulting information systems, as well as inter-organizational information systems in general are typically covered by such an analysis. Joint application development bundles development capabilities in an effort to reduce development costs. Typically specific know how and information is shared between the cooperating development partners. Hence, in case of competing developers, it is necessary to compare the benefits associated with reduced costs to possible disadvantages faced by disclosing information and know how. In this paper we will assume that information is shared via temporary joint. Note, that in our context collaboration may be characterized as being pre-competitive. It should not be mixed up with collusion which may be legally restricted or even forbidden. A formal model will be developed in the sequel. Techniques from game theory allow to solve the corresponding optimization problems. The model will be analyzed in a simple setting in order to be able to derive closed-form solutions which may be handled more conveniently. Hence, it seems natural to assume that a rational company maximizes its expected profits. The compensation of such managers is very often tied to profits. This fact, as well as possible opportunistic behavior and asymmetric information, suggest that managers behave more or less risk averse (Kao and Hughes, 1993). Consequently, expected utility of profits is maximized instead of expected profits.

3 REINFORCEMENT LEARNING

Learning from reinforcements has received substantial attention as a mechanism for robots and other computer systems to learn tasks without external supervision. The agent typically receives a positive payoff from the environment after it achieves a particular goal, or, even simpler, when a performed action gives good results. In the same way, it receives a negative (or null) payoff when the action (or set of actions) performed brings to a failure. By performing many actions overtime (trial and error technique), the agents can compute the expected values (EV) for each action. According to Sutton and Barto (1998) this paradigm turns values into behavioral patterns. Most RL algorithms are about coordination in multi agents systems, defined as the ability of two or more agents to jointly reach a consensus over which actions to perform in an environment. An algorithm derived from Q-Learning (Watkins, 1989) can be used. The EV for an action is updated every time an action is performed, according to Kapetanakis et al (2004):

$$EV_{t+1}(a) = EV_t(a) + \lambda(p - EV_t(a)) \qquad (1)$$

Where $0 < \lambda < 1$ is the learning rate and p is a payoff received every time action a is performed. This is particularly suitable for simulating multi stage games (Fudenberg and Levine 1998), in which agents must coordinate to get the highest possible aggregate payoff. Given a scenario with two agents (A and B), each of them endowed with two possible actions a_1, a_2 and b_1, b_2 respectively, the agents will get a payoff, based on a payoff matrix, according to the combination of performed actions.

3.1 Ego Biased Learning

While discussing the cognitive link among preferences and choices is definitely beyond the purpose of this work, it's important to notice that it's commonly accepted that the mentioned aspects are strictly linked among them. The link is actually bi-directional (Chen, 2008), meaning that human preferences influence choices, but in turn the performed actions (consequent to choices) can change original preferences. As stated in Sharot et al. (2009): "…*past preferences and present choices determine attitudes of preferring things and making decisions in the future about such pleasurable things as cars, expensive gifts, and vacation spots*".

Even if preferences can be modified according to the outcome of past actions (and this is well represented by the RL algorithms described before), humans can keep an emotional part driving them to prefer a certain action over another one, even when the latter has proven better than the former. Some of these can be simply wired into the DNA, or could have formed in many years and thus being hardly modifiable. A bias is defined as "*a particular tendency or inclination, esp. one that prevents unprejudiced consideration of a question; prejudice*" (www.dictionary.com).

That's the point behind learning: human aren't machines, able to analytically evaluate all the aspects of a problem and, above all, the payoff deriving from an action is filtered by their own perception bias. There's more than just a self-updating function for evaluating actions and in the following a formal reinforcement learning method is presented which keeps into consideration a possible bias towards a particular action, which, to some extents, make it preferable to another one that has analytically proven better through the trial and error period. Ego Biased Learning allows to keep this personal factor into consideration, when applying a RL paradigm to agents.

In the presented formulation, a dualistic action selection is considered, i.e.: $A(a_1, a_2)$. In our context, alternative actions represent two dyadic collaborative behaviour: the whole of strategic and tactic enterprise activities to establish a collaboration (a_2) or not (a_1). In the last case (a_1), enterprises carry on other competitive strategies non based on collaboration. By applying the formal reinforcement learning technique described in equation (1) an agent is able to have the expected value for the action it performed. Each agent is endowed with the RL technique. At this point, we can imagine two different categories of agents (α_1, α_2): one biased towards action a_1 and the other one biased towards action a_2. For each category, a constant is introduced $(0 < K_1, K_2 < 1)$, defining the expectation of each action, used to evaluate $\overline{EV(a_1)}$ and $\overline{EV(a_2)}$ which is the expected value of the action, corrected by the bias. For the category of agents biased towards action a_1 we have that:

$$\alpha_1: \begin{cases} \overline{EV(a_1)} = EV(a_1) + (|EV(a_1)| * K_1) \\ \overline{EV(a_2)} = EV(a_2) - (|EV(a_2)| * K_1) \end{cases} \quad (2)$$

In this way, K_1 represents the propensity for the first category of agents towards action a_1 and acts as a percentage increasing the analytically computed $EV(a_1)$ and decreasing $EV(a_2)$. At the same way, K_2 represents the propensity for the second category of agents towards action a_2 and acts on the expected value of the two possible actions as before. The constant K acts like a "friction" for the EV function; after calculating the objective $EV(a_i)$ it increments it of a percentage, if a_i is the action for which the agent has a positive bias, or decrements it, if a_i is the action for which the agent has a negative bias. In this way, the agent α_1 will perform action a_1 (instead of a_2) even if $EV(a_1) < EV(a_2)$, as long as $\overline{EV(a_1)}$ is not less than $\overline{EV(a_2)}$.

If $\overline{EV(a_1)} = \overline{EV(a_2)}$, by definition, the performed action will be the favorite one, i.e.: the one towards which the agent has a positive bias.

4 EXPERIMENTAL RESULTS

Some experiments were performed in order to test the basic EBL equations introduced in paragraph 3.1. The agents involved in the simulation can perform two possible actions, a_1 and a_2. The agents in the simulation randomly meet at each turn (one to one) and perform an action according to their EV. A payoff matrix is used, where p_1 is the payoff originated when both agents perform a_1 (both of enterprises don't want establish a collaboration), p_2 is the payoff given to the agents when one of them performs a_1 and the other one performs a_2 and so on (only one of them want – and try without success – establish a collaboration. Usually p_2 and p_3 are set at the same value, for coherency. If corss-strategies are the same and based on collaboration, pay off (p_4) is maximum. For each time-step in the simulation, the number of agents performing a_1 and a_2 are sampled and represented on a graph.

In the first experiment, a small bias towards action a_1 is introduced for fifty α_1 agents (K_1 = 0.1). Agents α_2 do not have a bias, but all start playing action a_2 which is the most favourable one, according to the payoff matrix; this will be different in the following experiments, where unbiased agents will start performing a random action. The results are quite interesting, and depicted in figure 1.

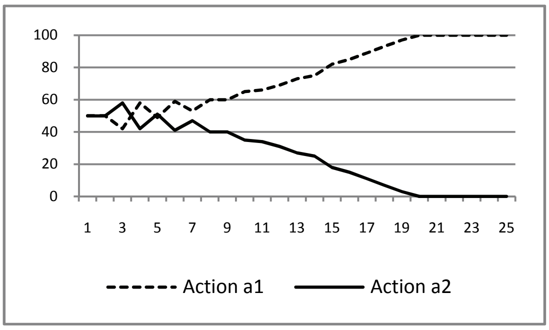

Figure 1: Experiment 1: biased Vs unbiased agents.

In this example it has been chosen to have all the rational agents (the straight line) starting from the favoured action (a_2) so to show that, even so, it is enough to have one half of the agents acting not completely rationally to make the system go towards the sub-optimal action (a_1). In fact, even if action a_2 is clearly favored by the payoff matrix (payoff 2 vs 1), after taking an initial lead in agents' preferences, all the population moves towards action a_1. This is due to the resilience of biased agents in changing their mind; doing this way, the other 50 non-biased agents find more and more partners performing action a_1, and thus, if they perform a_2 they get a negative payoff. In this way, in order to gain something, since they are not biased, they are forced to move towards the sub-optimal action a_1, preferred by the biased agents. In order to give a social explanation of this, we can think to the fact that often the wiser persons adapt themselves to the more obstinate ones, when they necessarily have to deal with them, even if the outcome is not the optimal one, just not to lose more. This is particular evident when the wiser persons are the minority, or, as in our case, in an equal number.

Till now the advantage of performing join action $a_2 + a_2$ over $a_1 + a_1$ was evident (payoff 2 vs 1) but not huge; in the next experiment, a new payoff matrix is used, in the joint action $a_2 + a_2$ is rewarded 3, instead of 2. The purpose is investigating how much the previous threshold would increase under these hypotheses. The empirical finding is 25/75, and the convergence is again extremely fast, and much similar to the previous experiment. Even a bigger advantage for the optimal action is soon nullified by the presence of just 25% biased agents, when penalty for miscoordination exists. This explains why sometimes suboptimal actions (or non-best products) become the most common. In real world, marketing could be able to bias a part of the population, and a good distribution or other politics for the suboptimal product/service could act as a penalty for unbiased players when interacting with biased ones.

5 CONCLUSIONS

While individual preferences are very important as a bias factor for learning and action selection, when dealing with social systems, in which many entities operate at the same time and are usually connected over a network, other factors should be kept into consideration, when dealing with learning. Ego biased learning is formally presented in the most simple case, in which only two categories of agents are involved, and only two actions are possible (collaboration or not). That's to show the basic equations and explore the results, when varying the parameters.

Some simulations are run, and the results are studied, showing how, even a small part of the population, with a negligible bias towards a particular action, can affect the convergence of the whole population. In particular, if miscoordination is punished (when cross-strategies are different), after few steps all the agents converge on the suboptimal action, which is the one preferred by the biased agents. With no penalty for miscoordination things are less radical, but once again many non-biased agents (even if not all of them) converge to the suboptimal action (non collaborative actions). This shows how personal biases are important in social systems, where agents must coordinate or interact.

If we look at things from a managerial/sociologic point of view, we have the following explanation.

The presented experiments show that few players potentially adverse to exchange information in a system, are enough for all the players to stop exchanging. This happens because the higher risk aversion of these operators brings all the others to the idea that carrying on collaborative strategies is a potential dispersion of resources. In fact, whenever a collaborative player crosses a non-collaborative one, they both evaluate the possible business, but after that the non-collaborative player denies it. From this, the penalty for miscoordination. A collaborative rational agent, after meeting some non-collaborative ones, changes her mind as well, since each time she loses some resources. Then, she becomes non-collaborative as well, unless she finds many collaborative players in a row. In other terms, to avoid a refusal after trying to collaborate, which is something that waste time and resources, also potentially collaborative agents will start to immediately refuse the possibility of a cooperation. By doing this, they won't gain as much as they would through collaboration, but they won't also risk to lose resources. The whole system thus settles on the sub-optimal equilibrium, in which no player collaborates.

In future works, general cases will be faced (more than two possible actions, different biases) in order to analyze the psychological drivers behind firms collaborations and additional experiments will be run.

REFERENCES

Chen M. K., 2008. Rationalization and Cognitive Dissonance: do Choices Affect or Reflect Preferences? Cowles Foundation Discussion Paper No. 1669

Clemons, E., Reddi, S. Row, 1993, M.: The Impact of Information Technology on the Organization of Economic Activity – The "Move to the Middle" Hypothesis. Journal of Management Information Systems 10, No. 2, pp. 9-35.

Fudenberg, D., and Levine, D. K. 1998. The Theory of Learning in Games. Cambridge, MA: MIT Press

Jin, J.: 1994. Information Sharing Through Sales Report. Journal of Industrial Economics 42, No. 3

Kao, J., Hughes, J. 1993. Note on Risk Aversion and Sharing of Firm-Specific Information in Duopolies, Journal of Industrial Economics 41, No. 1

Palfrey, T. R., 1982. Risk Advantages and Information Acquisition. Bell Journal of Economics 13, No. 1, pp. 219-224.

Powers R. and Shoham Y., 2005. New criteria and a new algorithm for learning in multi-agent systems. In Proceedings of NIPS.

Sharot T., De Martino B., Dolan R.J., 2009. How Choice Reveals and Shapes Expected Hedonic Outcome. The Journal of Neuroscience, 29(12):3760-3765

Sutton, R. S. and Barto A. G., 1998. Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA. A Bradford Book

Watkins, C. J. C. H., 1989. Learning from delayed rewards. PhD thesis, Psychology Department, Univ. of Cambridge.

Williamson, O., 1975. Markets and Hierarchies – Analysis and Antitrust Implications. New York.