Communication in Games

Timo Ehrig, Diana Garcia Lopez, Jolie Martin, Peng Peng, Tong Tong
CSSS, Beijing, 2008

I. Introduction and Motivation

Traditional game theory assumes that all players have full information about the interpersonal decision process and the opponent's incentives and abilities. It is assumed that players react "mechanical" to a given situation, and that alter knows that ego will react "mechanical" as well. In real-world situation, communication can change the mode of interaction between players. We intend to bring this idea into formal modeling.

Evolutionary game theory extends the framework of game theory allowing to consider the history of an interaction. Players try to guess what the opponents will do in the next situation based on experience in the history of the game.

We ask a simple question to investigate the effect of communication in a repeated interaction: What happens if one of the players sends a message to the other, about the action he will do, before the latter takes his guess? Such communications may change the expectation of the message receiver and the equilibrium of the game. An interesting question is how the receiver decides whether to trust the message, and how the sender decides whether to follow through.

Moreover, in time serial, if the players have "memory" they may predict the other's future action according to his behavior during the past and a trustworthy relationship between the players can emerge.

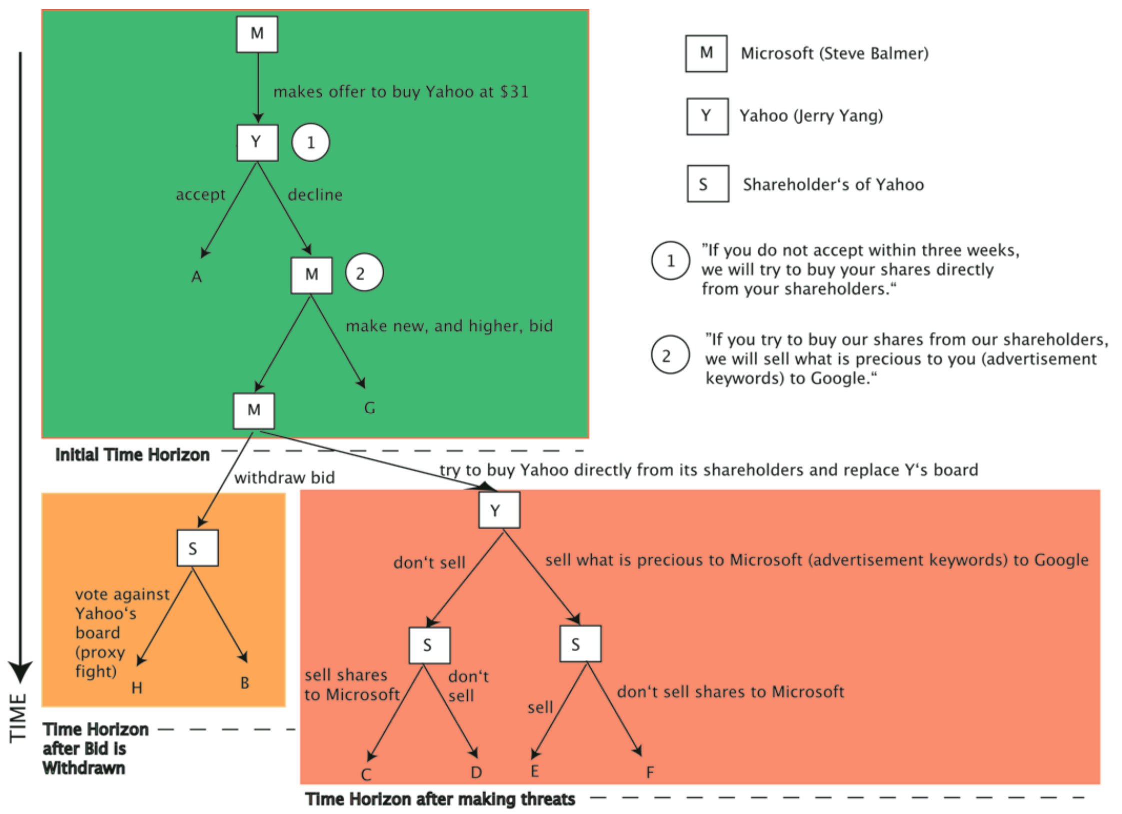
We first discuss a real case in the market, which shows the important role of communication; then introduce a time serial model of communication game with only two players and compare with the traditional game without communication; finally we give out a possible analytical result of this model.

II. The Microsoft Bid on Yahoo! Case

In game theory it is usually assumed that players have a shared understanding about the possible moves in their interaction, and shared knowledge about the pay-offs of outcomes for each party. In particular, it is most often assumed that the players know when the game will end, and how desirable such the possible end-points are to the players.

Real-world interactions are, however, characterized by contingencies. Each player can introduce possible new moves. The time horizon of a game is not given, but enacted in the course of interaction. Still, in many cases, real-world actors can be assumed to reason strategically, i.e. by playing though possible moves that seem likely to them. Communication, as we will show for the case of Microsoft's bid on Yahoo and explore in our general model, influences the anticipated likelihood of possible future moves.

We use real merging case, Microsoft's bid on Yahoo! in 2008, to illustrate properties of communication in real-world interaction. We will then proceed and model one of these properties in a formal model.



The above illustration models the interaction between Microsoft, Yahoo!, and Yahoo's shareholders in February – July 2008. It translates the interaction into a game tree in extensive form.

The green area indicates the initial situation of the bid: Microsoft made the offer, Yahoo can accept or decline, then Microsoft can make a new bid or withdraw. This created a game with outcomes A, G, and the possibility for Microsoft to withdraw the bid.

Our first observation is that the players *extended the possibilities of their interaction by means of communication*.

The "red area" of the interaction appeared by means of two threats: Steve Balmer threatened Jerry Yang to try to buy Yahoo directly from Yahoo's shareholders and to replace the board, if Yang wouldn't sell. Yang reacted and threatened Balmer that he would sell what is precious to Microsoft (advertisement keywords) to Google if

Balmer would really approach the shareholders directly.

We note that after the communications, the players could reason in form of a game again, and consider the possibilities given in the red area. However, the communication changed the structure of the game. The orange area in the game-set up indicates a third time horizon that was not yet mentioned in the early interaction between Microsoft and Yahoo, but that is a likely future horizon. This horizon might influence how much Yahoo desires that Microsoft withdraws the bid.

We take two observations from this real-world case:

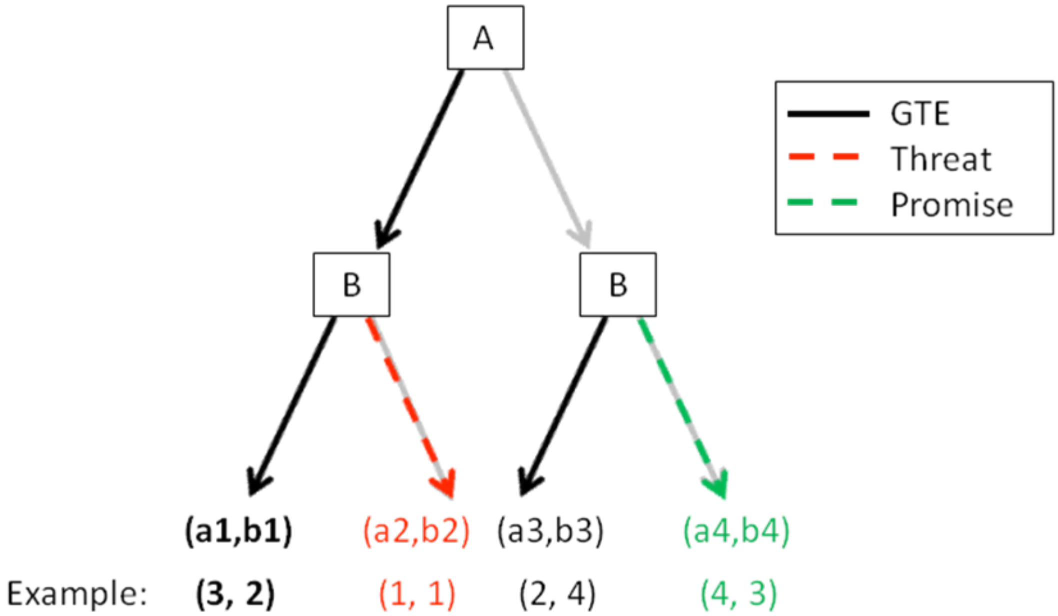
1. How the communication changes the anticipated likelihood of future moves depends on how much the opponent is believed to follow through the future action he committed to. It is to be discussed how communication should affect the assumed probabilities of the opponent's moves in the game tree.
2. The anticipation of future actions of the opponents also depends on the likelihood of *not yet mentioned, but possible, moves* that were used and were successful in similar interactions in the past. For instance, the orange time horizon, even though not explicitly addressed in Spring 2008 in the communications of the players, can be assumed to be already present in their strategic reasoning.

In our formal model, we focus on observation 1.: How does the history of the interaction between players influence alter's believes of ego following through his committed action?

III. Game setup and simulation results

Game Setup

In this sequential decision tree, Player A (first-mover) and Player B (second-mover) know their own payoffs for each outcome and the preference ordering over outcomes for the other player, so the game theoretic equilibrium (GTE) can be determined by backward induction. However, our model incorporates the possibility of pre-game communication by Player B of his intended action contingent on an action of Player A.



	Example Conditions for GTE=1, Threat=2	Example Conditions for GTE=1, Promise=4
Communication Goal: B wants A to choose opposite branch	b3 > b1	b4 > b1
Communication Need: B's stated action wouldn't maximize own payoff (isn't self-evident)	b1 > b2	b3 > b4
Communication Impact: A is better off choosing opposite branch if he trusts B, not if he doesn't	a1 > a3 > a2	a4 > a1 > a3

Model Parameters

- n1 rounds of n2 games each
- payoffs a1, b1, a2, b2, a3, b3, a4, b4 drawn randomly ~U(0,1) for each game
- w = trustworthiness of Player B (actual tendency to follow through on threats or promises) drawn randomly ~U(0,1) for each round
- t = trust of Player A in Player B (belief that he will follow through on threats or promises) drawn randomly ~U(0,1) at the beginning of each round, then updated after each game

Player Strategies

- Player B sends a message (threat and/or promise) if payoff conditions are met
- Player A maximizes expected value of own payoffs given current trust t
  - Example: (1-t)*a1 + t*a2 <?> (1-t)*a3 + t*a4
- If Player B is tested, he follows through on threat or promise with probability w
- If Player B was tested, Player A updates his trust as a weighted average of old belief and latest observation of Player B's behavior: $t_{new} = (1-c)*t_{old} + c*obs$
  - *obs* = {1 if Player B followed through on threat or promise; 0 otherwise}
  - c = 1 / (1 + *# tests so far*)

Simulation Results

We examined the outcomes of a simulation with n1 = 100 rounds of n2 = 1000 games each.

First, to explore how well Player A was able to learn the trustworthiness of Player B based simply on observation of his tendency to follow through, we measured the average ratio t/w in each of the sequential games, averaged across rounds. As shown in Figure 1, this ratio initially has a great deal of variability, with Player A's trust in Player B deviating a great deal from Player B's trust trustworthiness, but eventually converges to 1 as Player A has sufficient observations to estimate Player B's trustworthiness parameter.

Next, we looked at the cumulative payoffs for all games of each player as a function of Player B's true trustworthiness for that round. Figure 2 demonstrates that

Player B receives higher aggregate payoffs when he is trustworthier, while the opposite is true for Player A. The reason for this effect is that for random payoffs, a threat scenario where Player B is much more likely to occur than a promise scenario, and there is an asymmetrical benefit to following through in each. For a threat, if Player B proves himself credible to follow through, he can coerce Player A to do as he wishes even though this outcome is worse for Player A than the GTE outcome. On the other hand, only for promises would both players be able to benefit from Player B's credibility, but since they occur more rarely, Player A has few opportunities to benefit from Player B following through.

A similar pattern is observed in Figures 3 and 4 where we compare Player A and Player B's payoffs, respectively, to the payoffs they would have received through realizing the GTE outcome. Although the roles are not symmetric in the sequential game we describe, it is clear that Player A performs worse than he would have otherwise whereas Player B performs better given the particular strategic modification we have modeled with the potential for Player B to send messages about his future intentions. It could be argued that in the real world Player B would fail to follow through on these stated intentions since he actually could gain higher payoffs in each game by defecting, yet we have attempted to demonstrate one mechanism by which a game participant might build up his credibility by following through and, thus, realize long run gains.

Figure 1

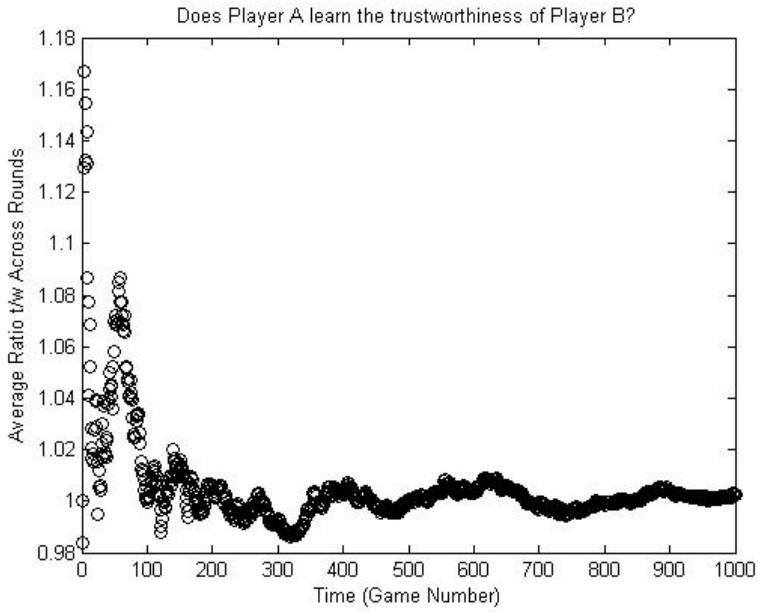


Figure 2

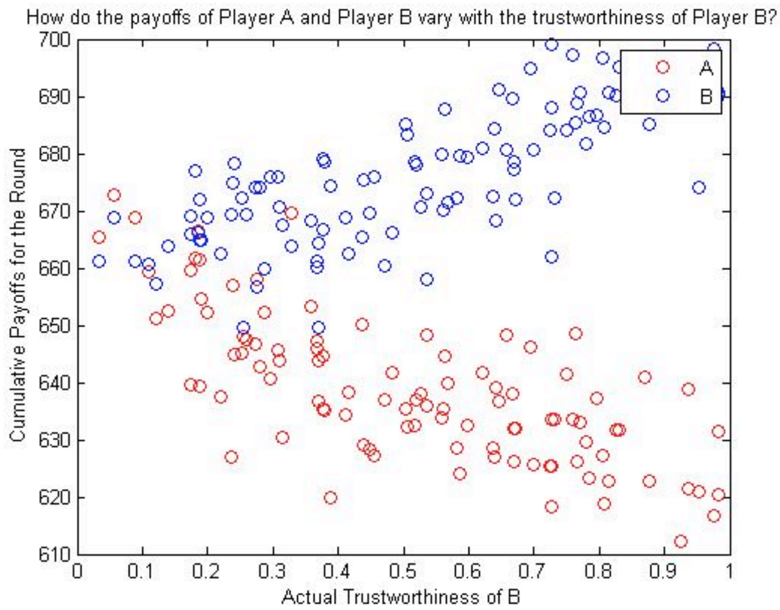


Figure 3

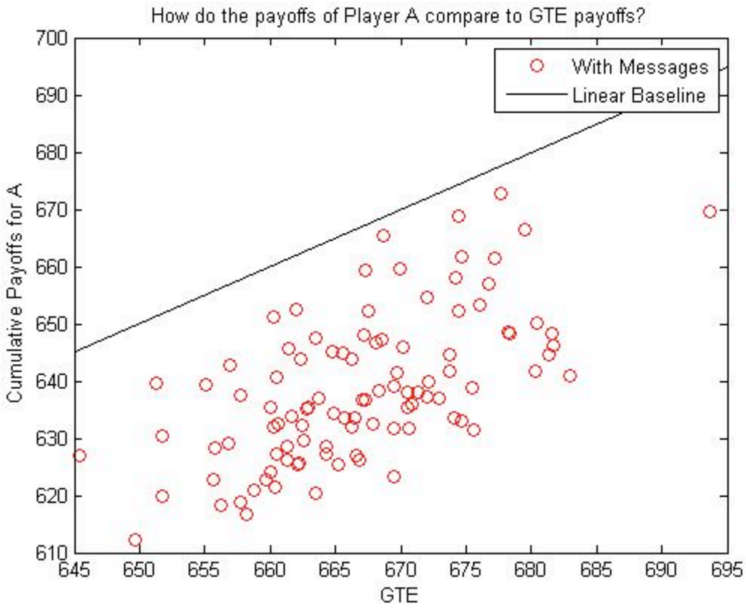
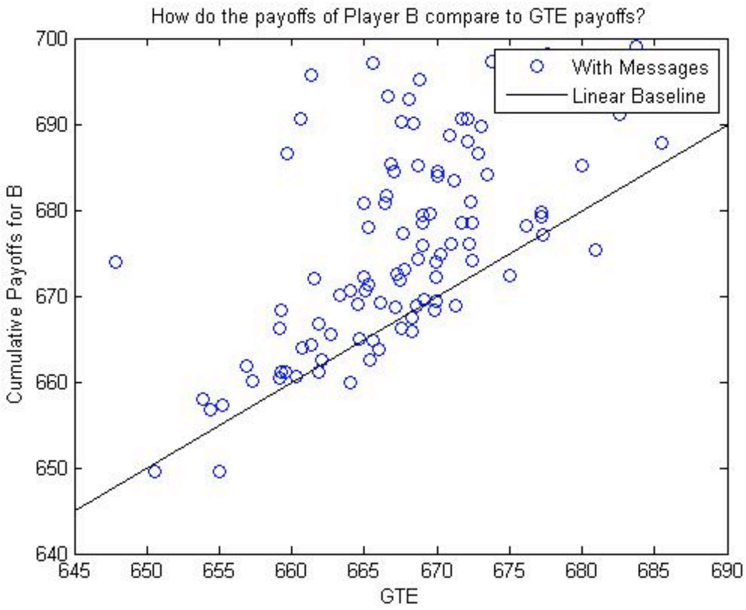


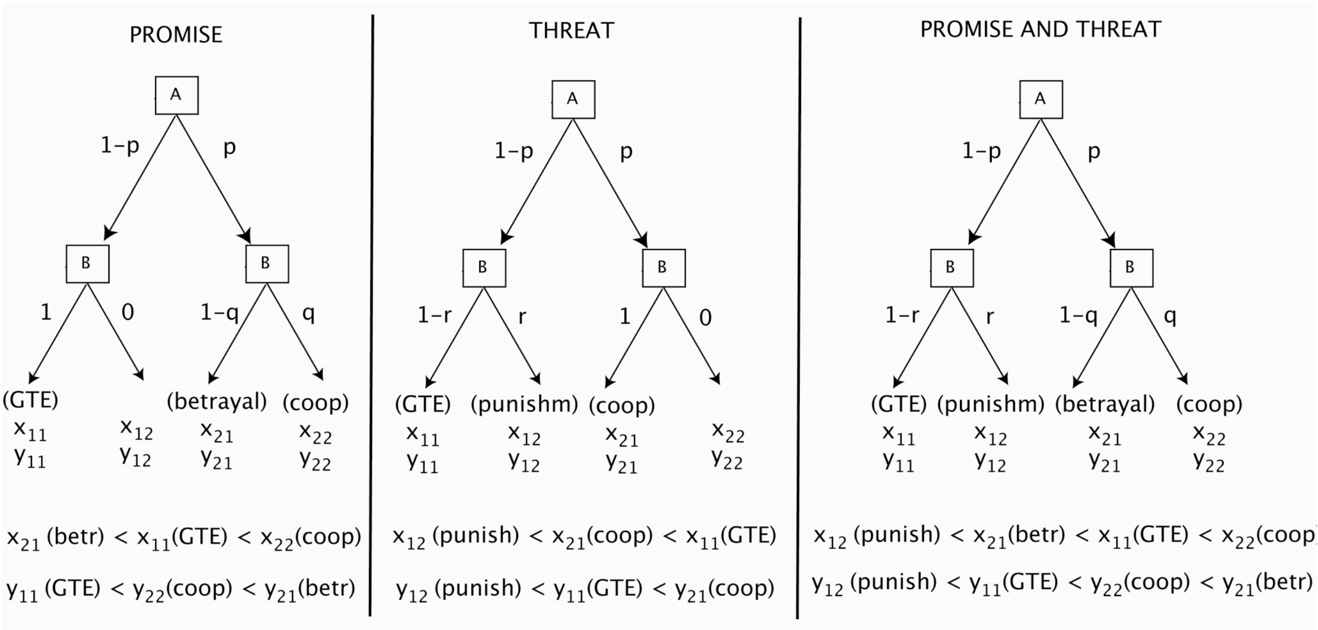
Figure 4



IV. Optimum strategy analysis

Let's take a small part of this general simulation model and analyse it, passing from a dynamic to a static view of the problem: given a certain promise/threat game tree, what would the optimum strategies of the two players be? In other words, what would the steady state, or long time limit, of a learning game be (after a repeated interaction)?

- x, y: payoffs for A ($x = \{x_{ab}\} = \{x_{11}, x_{12}, x_{21}, x_{22}\}$) and B ($y = \{y_{ab}\}$).
- p(x, q(y)), q(y): weights on branches ~ strategies of A and B.

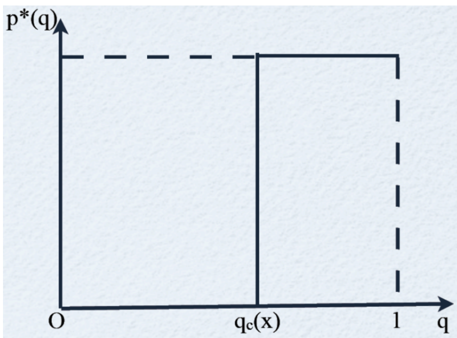


Promise situation

To gain understanding we first consider a pure promise situation. Player A wishes to maximise his expected payoff, which is (given the promise tree):

$$E[x] = \sum_{a,b} x_{ab}\, Pr(a,b) = \sum_{a,b} x_{ab}\, Pr(b|a)\, Pr(a) = (1-p)x_{11} + p(1-q)x_{21} + p\, q\, x_{22} =$$
$$= x_{11} + p((x_{22} - x_{21})q - (x_{11} - x_{21})) \equiv x_{11} + p(d_2\, q - d_1)$$

A's optimum trust in B will be



$p* = max_p(E[x])$

so we find

$$\frac{dE}{dp} = d_2\, q - d_1 \begin{cases} > 0 & \Rightarrow p* = 1 \\ = 0 & \Rightarrow p* \text{ undefined} \\ < 0 & \Rightarrow p* = 0 \end{cases}$$

i.e. the optimum trust of A in B is a step-function:

$$p^*(x,r) = max_p\{E[x](x,p,r)\}$$

$$= max_p\{\sum_{a,b} x_{ab}\, Pr(b|a)\, Pr(a))\} = \Theta(q - q_c(x))$$

the jump being at

$$q_c(x) = \frac{x_{11} - x_{21}}{x_{22} - x_{21}}$$

which can be undestood as:

$$\frac{(x_{22}-x_{11})\,q_c}{(x_{22}-x_{21})(1-q_c)} = \frac{(\text{gain if coop})\cdot Pr(\text{coop})}{(\text{loss if betr})\cdot Pr(\text{betr})} = 1$$

In other words, player A must compare his expected gain, should B keep his word (cooperate), to the expected loss he would incur in, should B betray him (not follow through), where "gain" and "loss" are interpreted as deviations from the GTE equilibrium payoff.

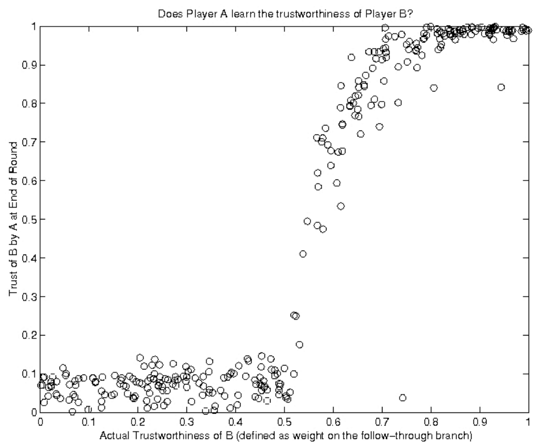
Modified dynamic simulation (promise)

How does this relate to our simulation? Let's go back to it and modify it: only consider promises and, instead of making B follow through with a certain fixed probability, let him consider his own payoffs, and follow through with probability q_{eff}:

$$q_{\text{eff}} = Pr(y_{\text{coop}}\cdot q_0 \;>\; y_{\text{betray}}\cdot(1-q_0)\,)$$

Remember that y_{coop} and y_{betray} are random variables drawn uniformly, U[0,1], and satisfying $y_{coop} > y_{betray}$; and that q_0 is also U[0,1], and kept constant through each round. The simulation just reflects A's learning q_{eff} as a frequentist. The following plot shows the value of q_{eff} at the final timesteps as a function of B's actual trustworthiness q_0 (one dot per round).

$$q_{\text{eff}}(q_0)\ \text{ vs. }\ q_0$$



Threat situation

A similar analysis holds for the pure threat situation. A's optimum belief in B's threat is

$$p^*(x, r) = max_p \{E[x](x, p, r)\} = \Theta(r - r_c(x))$$

with the jump of the step function being at

$$r_c(x) = \frac{x_{11} - x_{21}}{x_{11} - x_{12}}$$

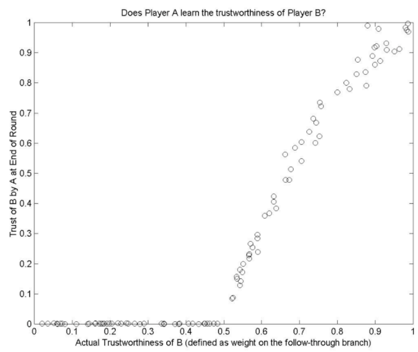
which satisfies

$$\frac{(x_{11} - x_{12})\, r_c}{x_{11} - x_{21}} = \frac{(\text{loss if punishm}) \cdot Pr(\text{punishm})}{(\text{loss by coop})} = 1$$

In a threat game, player A must compare his expected loss, should B keep his word and carry out his threat (punish him for not cooperating), to the expected loss he would suffer through listening to B and choosing the other path (cooperation), where again these "losses" are deviations from the GTE payoff. Here we see the substantial difference between the promise and threat situations: a threat situation will always imply A getting a lower payoff than the GTE one (as seen in a previous graph).

Modified dynamic simulation (threat)

Once again we run pure threat simulations with a modified follow-through rule for B (taking into account his own payoffs). The end of the learning game now gives, for $r_{eff}(r_0)$ vs. r_0 :



Player b's optimum strategy

Up to now, B's behaviour was fixed, and A tried to optimize his own response. Let's imagine now that B has access to A's payoffs and can thus calculate the step functions we have just found. In order to keep A on the "believing" side of the threshold, while at the same time getting away with as much cheating as he can, B should set q just above the threshold: cooperate as little as necessary to maintain trust, i.e. to maintain the usefulness of the communication.

Promise and threat situation

A simultaneous promise and threat situation also gives a step function for A's best strategy. Considering B's behaviour characterised by 2 variables, q_1 and q_2, it turns out that B's optimum strategy is to always carry out his threat and to betray his promise as much as he can get away with (the threshold being slightly modified by the presence of the threat). The effect of adding the possibility of a threat is thus to reinforce the "arguments" that B uses to make A follow the "cooperation" branch of the tree.

Conclusion of the analytical attempt

By simplifying the problem we fell back to doing some basic classical tree analysis to find the optimum strategies. When wishing to compare the simple results to a modified version of the computer simulation, we find something which also shows a threshold but actually only reflects the combined effect of B's decision mechanism and the random payoff structure on the otherwise straighforward learning game, so no comparison is possible (a hasty look would want to see p* vs. q in the simulation results graph, but it only shows q_{eff} vs. q_0, given that p->q_{eff} as A learns).

The long time limit of the original general simulation in which payoffs are random (and all three promise/ threat/ promise&threat situations may arise) would correspond to a weighted average of the three types of games (weighted by the probability that the payoffs drawn meet the requirements). The fun possibly lies, as we said, in making B learn too, and in the consideration of different dynamics of learning, of shorter time horizons.

V. Conclusion

We conclude as following from the analysis:

- Communication of intentions will fundamentally change the results of repeated interaction games.
- Trust and trustworthiness can survive through mutual reinforcement.

Although we have analyzed a simplified model in order to deal with the complexity from the real world, still it is hard to find a general solution of the communication in repeated game, let alone the real negotiations in business. However, it is meaningful to attempt to explain the reality.

Future work includes investigating a symmetric setting, i.e. to allow both players to make threats and promises.

Furthermore, we hope to find models that can account for observation 2. of our real-world example: How can we account for the fact that players anticipate opponent's moves that seem likely from past interactions, that are not included in the given game, but possible contingencies?

VI. Acknowledgements

This collaboration was initiated at the 2008 China Complex Systems Summer School, co-sponsored by the Santa Fe Institute and the Institute of Theoretical Physics of the Chinese Academy of Sciences. The research and education programs of SFI are supported by core funding from the U.S. National Science Foundation and by gifts and grants from individuals, corporations, other foundations, and members of the Institute's Business Network for Complex Systems Research. SFI

received direct support for the 2008 China CSSS from the US NSF under award OSIE-0623953.

We would like to specifically thank professors Susanne Still and Will Tracy, and everyone who took part in the summer school.

VI. References

Farrell, Joseph; Rabin, Matthew (1996); Cheap Talk, Journal of Economic Perspectives, Vol. 10, No. 3, pp. 103-118

Rubinstein, Ariel, (1998) Modeling Bounded Rationality, MIT Press

Selten, Reinhard, (1978) The Chain Store Paradox, Theory and Decision 9, 127-159.