Game Theory (continued)*

There are at least two interesting ways in which noise can be introduced into the play of a repeated game: (1) transmission noise; and (2) misperception. In the prisoner's dilemma, for instance, a decision pair (C_i, C_j) can be erroneously transmitted as any of the other three possible outcomes with some small probability (say ϵ), before it is acted upon by the players for the next turn. So, the "intent" (C_i, C_j) is transmitted into an observation

(C_i, C_j)	with probability $(1-3\epsilon)$
(C_i, D_j)	with probability ϵ
(D_i, C_j)	with probability ϵ
(D_i, D_j)	with probability ϵ

Of course, since transmission noise applies to any intent, any of the four possible decision pairs gives rise to any of the four possible observations, thus defining a Markov chain with (transmission noise) transition matrix $N(\epsilon)$

$$N(\epsilon) = \begin{array}{c} \\ (C_i, C_j) \\ (C_i, D_j) \\ (D_i, C_j) \\ (D_i, D_j) \end{array} \begin{array}{c} \begin{array}{cccc} (C_i, C_j) & (C_i, D_j) & (D_i, C_j) & (D_i, D_j) \end{array} \\ \begin{pmatrix} 1-3\epsilon & \epsilon & \epsilon & \epsilon \\ \epsilon & 1-3\epsilon & \epsilon & \epsilon \\ \epsilon & \epsilon & 1-3\epsilon & \epsilon \\ \epsilon & \epsilon & \epsilon & 1-3\epsilon \end{pmatrix} \end{array}$$

In the case of misperception each side may misperceive the other side's intent into its opposite with probability ϵ. This would yield the misperception transition matrix $M(\epsilon)$

$$M(\epsilon) = \begin{array}{c} \\ (C_i, C_j) \\ (C_i, D_j) \\ (D_i, C_j) \\ (D_i, D_j) \end{array} \begin{array}{c} \begin{array}{cccc} (C_i, C_j) & (C_i, D_j) & (D_i, C_j) & (D_i, D_j) \end{array} \\ \begin{pmatrix} (1-\epsilon)^2 & \epsilon(1-\epsilon) & \epsilon(1-\epsilon) & \epsilon^2 \\ \epsilon(1-\epsilon) & (1-\epsilon)^2 & \epsilon^2 & \epsilon(1-\epsilon) \\ \epsilon(1-\epsilon) & \epsilon^2 & (1-\epsilon)^2 & \epsilon(1-\epsilon) \\ \epsilon^2 & \epsilon(1-\epsilon) & \epsilon(1-\epsilon) & (1-\epsilon)^2 \end{pmatrix} \end{array}$$

Suppose that the two players of a repeated prisoner's dilemma play TFT within a noisy environment. Without the noise, the TFT players also define a Markov chain on the set of four possible states $S = \{(C_i, C_j), \quad (C_i, D_j), \quad (D_i, C_j), \quad (D_i, D_j)\}$ with transition matrix T

$$T = \begin{array}{c} \\ (C_i, C_j) \\ (C_i, D_j) \\ (D_i, C_j) \\ (D_i, D_j) \end{array} \begin{array}{c} \begin{array}{cccc} (C_i, C_j) & (C_i, D_j) & (D_i, C_j) & (D_i, D_j) \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

---

*This section is copyrighted (c) by Jean-Pierre P. Langlois, 1999.

Of course, (C_i, C_j) is a closed recurrent class in this Markov chain as expected. But when noise (say transmission) is introduced, the transition matrix becomes $TN(\epsilon)$ as follows

$$TN(\epsilon) = \begin{array}{c} \\ (C_i,C_j) \\ (C_i,D_j) \\ (D_i,C_j) \\ (D_i,D_j) \end{array} \begin{array}{c} \begin{array}{cccc} (C_i,C_j) & (C_i,D_j) & (D_i,C_j) & (D_i,D_j) \end{array} \\ \begin{pmatrix} 1-3\epsilon & \epsilon & \epsilon & \epsilon \\ \epsilon & \epsilon & 1-3\epsilon & \epsilon \\ \epsilon & 1-3\epsilon & \epsilon & \epsilon \\ \epsilon & \epsilon & \epsilon & 1-3\epsilon \end{pmatrix} \end{array}$$

Because this matrix is positive, the whole space forms a single class (the chain is irreducible) and there exists a *single* stationary distribution $\pi(\epsilon) = \pi(\epsilon)TN(\epsilon)$. Let

$$\pi(\epsilon) = (\pi_1, \pi_2, \pi_3, \pi_4) = \pi(\epsilon)TN(\epsilon)$$

A simple calculation shows that $\pi_1 = \pi_2 = \pi_3 = \pi_4 = \frac{1}{4}$ *independently* of $\epsilon > 0$. So, the TFT players in a noisy environment find themselves with a limit average payoff (by the ergodic theorem) of

$$\tfrac{1}{4}U_i(C_i, C_j) + \tfrac{1}{4}U_i(C_i, D_j)\tfrac{1}{4}U_i(D_i, C_j)\tfrac{1}{4}U_i(D_i, DC_j) = -\tfrac{1}{2}$$

It is easy to imagine that this can be improved upon. Indeed, TFT does not do well in noisy tournaments. TF2T in fact does substantially better.

It has been found *experimentally* that "generous" Tit-for-Tat (GTFT) is among the best performers in such noisy tournaments. In GTFT, one retaliates for an opponent's defection with only about 10% probability. This margin of tolerance allows the absorpsion of the noise without the tendency to be exploited (say by the 50% random defector) that TF2T has. However, there is no real theoretical underpinning for GTFT. In what follows, I will present a recent approach to that problem.

Consider a generous "Markov" strategy that always reciprocates cooperation and retaliates for defection with less than full probability
$p(D_i|C_iD_j) = \alpha_i$, and $p(D_i|D_iD_j) = \beta_i$, with $0 \leq \alpha_i < 1$ and $0 \leq \beta_i < 1$.

Two such players determine a Markov chain with transition matrix

$$T = \begin{array}{c} \\ (C_i,C_j) \\ (C_i,D_j) \\ (D_i,C_j) \\ (D_i,D_j) \end{array} \begin{array}{c} \begin{array}{cccc} (C_i,C_j) & (C_i,D_j) & (D_i,C_j) & (D_i,D_j) \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1-\alpha_i & 0 & \alpha_i & 0 \\ 1-\alpha_j & \alpha_j & 0 & 0 \\ (1-\beta_i)(1-\beta_j) & (1-\beta_i)\beta_j & (1-\beta_j)\beta_i & \beta_i\beta_j \end{pmatrix} \end{array}$$

Of course, such a strategy pair will maintain cooperation with certainty in the absence of noise just as TFT or GTFT do, and even if play does not start at (C_i, C_j). When noise is present (say transmission noise again) however, the effects of intent are modified according to $TN(\epsilon)$. As this happens, the stationary distribution moves away from $\pi(0) = (1,0,0,0)$ to some $\pi(\epsilon)$, as ϵ increases from 0. Since $N(\epsilon)$ is differentiable in ϵ, we may differentiate $\pi(\epsilon) = \pi(\epsilon)TN(\epsilon)$ to obtain

$$\pi'(\epsilon) = \pi'(\epsilon)TN(\epsilon) + \pi(\epsilon)TN'(\epsilon)$$

and evaluate at $\epsilon = 0$. Then, we can approximate $\pi(\epsilon)$ by $\pi(0) + \epsilon\pi'(0)$ (with a term in ϵ^2 neglected). Since the payoff effects of $\pi(0)$ are already known and common to all reciprocating Markov strategies, the effect of noise, as mitigated by their generosity, will be found (up to a negligible term in ϵ^2) in the term $\pi'(0)$. Since $N(0) = I$ and since

$$N'(\epsilon) = \begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix}$$

we have at $\epsilon = 0$, noting that $TN(0) = TI = T$, and that $\pi(0)T = \pi(0)$:

$$\pi'(0)(I - T) = \pi(0)N'(0)$$

or, if we let $\pi'(0) = (p_1, p_2, p_3, p_4)$ and omitting p_1 since $p_1 + p_2 + p_3 + p_4 = 0$:

$$(p_2, p_3, p_4)\begin{pmatrix} 1 & -\alpha_i & 0 \\ -\alpha_j & 1 & 0 \\ -\beta_j(1-\beta_i) & -\beta_i(1-\beta_j) & 1-\beta_i\beta_j \end{pmatrix} = (1,1,1)$$

This immediately yields $p_4 = \frac{1}{1-\beta_i\beta_j}$ and

$$p_2 = \frac{(1+\beta_j-2\beta_i\beta_j)+\alpha_j(1+\beta_i-2\beta_i\beta_j)}{(1-\beta_i\beta_j)(1-\alpha_i\alpha_j)}$$

$$p_3 = \frac{(1+\beta_i-2\beta_i\beta_j)+\alpha_i(1+\beta j-2\beta_i\beta_j)}{(1-\beta_i\beta_j)(1-\alpha_i\alpha_j)}$$

Now, the effects on i's limit-average payoffs are proportional (by ϵ) to

$$w_i = (0)p_1 + (-2)p_2 + (+1)p_3 + (-1)p_4$$

So, in order to optimize w_i with respect to α_i (and β_i) we differentiate to obtain (after simplification):

$$\frac{\partial w_i}{\partial \alpha_i} = -2\frac{\partial p_2}{\partial \alpha_i} + \frac{\partial p_3}{\partial \alpha_i} = \frac{(1-2\alpha_j)\left[(1+\beta_i-2\beta_i\beta_j)+\alpha_j(1+\beta j-2\beta_i\beta_j)\right]}{(1-\beta_i\beta_j)(1-\alpha_i\alpha_j)^2}$$

The quantity between brackets in the numerator is always strictly positive, and so is the denominator. So, $\frac{\partial w_i}{\partial \alpha_i}$ has the sign (and zero) of $(1 - 2\alpha_j)$. Clearly, for $\alpha_j < \frac{1}{2}$, α_i=1 is best and for $\alpha_j > \frac{1}{2}$, $\alpha_i = 0$ is best. But of course, the situation is symmetric (for α_j given α_i. The only way two such generous players will do equally best against each other is when $\alpha_i = \alpha_j = \frac{1}{2}$. A similar condition can be worked out for β_i and β_j. But it already appears that a whopping 50% generosity is called for in this noisy case. Indeed, tournament simulations indicate that this highly generous strategy survives better than any other (generous reciprocating Markov). This theory can be generalized to other payoff coefficients and to misperception noise with very similar results.