Blue, Red and Green
An Arms-Control Dilemma

Answers

WTFT: If -say- Blue unilaterally defects while Red and Green cooperate, neither of the latter retaliate at the next turn. So, Blue's payoff E_U^B of unilateral defection reads

$$E_U^B = 3 + dE_{CCC}^B = 3$$

since $E_{CCC}^B = \frac{0}{1-d} = 0$ is the expected payoff of always cooperating by all. Clearly, $E_U^B > E_{CCC}^B$ so that WTFT is not even a Nash equilibrium.

STFT: As Prof. Sanctions remarked, if one side -say Blue- defects, the other two will defect at the next turn (while Blue will return to cooperation). But at the next-to-next turn, since both Red and Green will have defected, all three of them will defect, and this will then go on forever. In other words, we will have the sequence of states:

$$DCC \to CDD \to DDD \to DDD \to ....$$

Clearly, $E_{DDD}^B = \frac{-5}{1-d}$, $E_{CDD}^B = -4 - \frac{5d}{1-d} = \frac{-4-d}{1-d}$ and the expected payoff E_U^B of unilateral defection by Blue reads

$$E_U^B = 3 - \frac{d(4+d)}{1-d} = \frac{3-7d-d^2}{1-d} < 0 = E_{CCC}^B$$

provided $3 - 7d - d^2 < 0$, which follows from $d > \frac{\sqrt{61}-7}{2} \simeq 0.4051$. So, under this condition STFT is a Nash equilibrium.

But Prof. Sanctions is obsessed with *perfect* equilibria. This requires that the above sequence of plays remains optimal for all concerned. Suppose, for instance, that we are in state DDD. Blue, for instance, could consider a unilateral deviation U from the expected Dfct at that state. This means that Blue would cooperate despite an expected return to DDD thereafter. This yields

$$E_U^B = -4 + dE_{DDD}^B = \frac{-4-d}{1-d} > \frac{-5}{1-d} = E_{DDD}^B$$

for any $d < 1$ (a requirement for a discount factor). So, Dfct is not optimal for Blue (and any other player) at DDD. STFT is therefore not perfect (but who is?).

GT: We already obtained the expected payoff at DDD. In GT the calculus for unilateral deviation is simpler:

$$E_U^B = 3 + dE_{DDD}^B = 3 - \frac{5d}{1-d} = \frac{3-8d}{1-d} < 0 = E_{CCC}^B$$

provided $d > \frac{3}{8}$. Again this is a Nash equilibrium. But again, a unilateral cooperate from DDD is best and the scheme is not perfect.

CTFT: The expected payoff E^B_{BG} to Blue of being guilty (being at state BG) is:

$$E^B_{BG} = -4 + dE^B_{CO} = -4$$

where CO denotes the state where no-one is guitly and everyone is expected to cooperate for a discounted value $E^B_{CO} = 0$. A unilateral Dfct (U) by Blue therefore yields

$$E^B_U = 3 - 4d < E^B_{CO} = 0$$

as long as $d > \frac{3}{4}$. Clearly, it is best for anyone to cooperate at CO when there is enough concern for the future. The two remaining issues are whether (a) it is best for Red and Green to retaliate at BG; and (b) it is best for Blue to comply with the punishment.

(a) Failing to punish Blue for Red (or Green) would yield:

$$E^R_U = -2 + dE^B_{CO} = -2 < 1 + dE^B_{CO} = 1$$

less than what Red gets by implementing the punishment.

(b) Failing to comply with the punishment for Blue would yield, since he would remain guilty:

$$E^B_U = -5 + dE^B_{BG} = -5 - 4d < -4 = E^B_{BG}$$

So, Blue is better off complying. So, CTFT is a perfect equilibrium (as Prof. Sanctions likes them).

PRC: We first compute (for Blue or anyone) the expected payoff from reaching the state DF (when p is the probability of remaining in state DF for one more turn):

$$E^B_{DF} = -5 + d\big(pE_{DF} + (1-p)E_{CO}\big) = \frac{-5}{1-dp}$$

while the expected payoff from multilateral cooperation is as usual $E^B_{CO} = 0$. For PRC to be a Nash equilibrium, we merely need that the payoff E^B_U of unilateral defection from CO be negative:

$$E^B_U = 3 + \frac{-5d}{1-dp} = \frac{3-d(5+3p)}{1-d} < 0$$

or $\quad d(5+3p) > 3$ (this requires $d > \frac{3}{8}$). But there are two states in this scheme: CO and DF. For the scheme to be a perfect equilibrium we need that it provides a best reply for each player (such as Blue) at DF. Here, Prof. Sanctions suggestion comes handy: a unilateral deviation U (a Coop instead of Dfct) at DF will not bring a *probabilistic* but a *certain* return to DF (one might want to distinguish one more state UDF where at least one side fails to implement punishment). So:

$$E^B_{UDF} = -4 + dE_{DF} = -4 - \frac{5d}{1-dp} = \frac{-4+d(4p-5)}{1-dp} \leq \frac{-5}{1-dp} = E^B_{DF}$$

is possible if: $p \leq \frac{5}{4} - \frac{1}{4d}$, a relation that requires $d > \frac{1}{5}$. So, PRC is a perfect equilibrium for a wide range of discount factors.

Interestingly, Prof. Sanctions is particularly interested in the highest possible p because of his fee (since it means he remains in business) but admits that the lowest possible p is the best hope for peace. He fears he has a conflict of interest.

The Ultra Nationalist Issue: Once Green has decided to defect forever, Blue and Red face a Chicken game with payoffs

Blue \ Red	Coop	Dfct
Coop	U= -2 U= -2	U= -4 U= 1
Dfct	U= 1 U= -4	U= -5 U= -5

Prof. Sanctions pessimism is well justified: Blue and Red might engineer cooperation between themselves but that will leave them vulnerable to green's exploitation. "This is a Free-Rider problem, we need more technology..." Sanctions is quoted as saying before leaving the office in disgust.