Chapter Five: Dynamic Games*

As mentioned earlier, the main conceptual difference between static and dynamic games is that the former has a preset finite number of turns while the latter can potentially last forever and ends only with a decision by a player or by chance. We will discuss three main kinds of dynamic games according to structure: (1) games defined on graphs, (2) repeated normal-form games, and (3) repeated continuous games.

5.1 Some General Principles

As mentioned in Chapter One, dynamic games raise two further issues about the players:

1. How do they remember the past?
2. How do they appraise the future?

The "history" of a dynamic game is simply a record of what all the players did at all prior turns of the game. If we denote by $\boldsymbol{\mathcal{A}}$ the (constant) action space of all the players and by $\boldsymbol{\mathcal{H}}$ the set of all possible histories of any length it is then the reunion:

$$\boldsymbol{\mathcal{H}} = \bigcup_{n=1}^{\infty} \left(\prod_{i=1}^{n} \boldsymbol{\mathcal{A}} \right)$$

If the game can last forever, this is an awfully large set of possible pasts for an average player to remember. So, in practice, players only remember finitely many possible developments called the "states" of the (repeated) game. In theory, strategies must stipulate how to respond to all possible hitories. But practically, players will identify some characteristics that can be shared by numerous possible histories. For instance, a player might only want to pay attention to what was done at the very last turn, or in the last couple of turns, or whether one particular move was ever used by the others and how frequently. This means that $\boldsymbol{\mathcal{H}}$ is *partitionned* into a set $\boldsymbol{\mathcal{S}} = \{\mathcal{S}_k\}_{k \in \mathbb{N}}$ of "states" $\mathcal{S}_k$ of the game that are shared by all players.[1]

In order for strategies to use current states $\mathcal{S}_k$ rather than the entire history $\boldsymbol{\mathcal{H}}$ in the description a player's reactions, it is also necessary to stipulate how states evolve from turn to turn. This requires a "transition" rule $\mathcal{T}$:

$$\mathcal{T} : \boldsymbol{\mathcal{S}} \times \boldsymbol{\mathcal{A}} \to \boldsymbol{\mathcal{S}} \qquad (1)$$

---

[1]A partition is a reunion, finite or infinite, of disjoint sets that, together make up the whole set. This means that $\boldsymbol{\mathcal{H}} = \bigcup_k \mathcal{S}_k$ and that, if $k \neq l$ then $\mathcal{S}_k \cap \mathcal{S}_l = \emptyset$. The condition that all players "share" the same partition is not a restriction. In fact, one player may contemplate a partition different from another. For instance, one player may consider only whether the *other* player chose Cooperate (C) or Defect (D) at the last turn, in a repeated Prisoner's Dilemma. But this translates into a "common" partition into four states $\{(\mathrm{C},\mathrm{C}), (\mathrm{D},\mathrm{C}), (\mathrm{C},\mathrm{D}), (\mathrm{D},\mathrm{D})\}$ for the entire history $\boldsymbol{\mathcal{H}}$.

where $\boldsymbol{\mathcal{A}}$ is the "action space", meaning the set of possible decisions by all players at any one turn. So, if $\mathcal{S}_t \in \boldsymbol{\mathcal{S}}$ is the state at turn t and a set of choices a_t is made by the players at turn t, then the next state is obtained through the transition rule by:

$$\mathcal{S}_{t+1} = \mathcal{T}(\mathcal{S}_t, a_t) \tag{2}$$

In practice, $\mathcal{T}$ formalizes the players' *interpretation* of history. The simplest example of such a transition rule is given by the two-states-memory repeated Prisoner's Dilemma of Chapter One: joint cooperation from the Cooperated state leads back to the same Cooperated state. Anything else leads to the Defected state forever. The players, therefore, interpret the past in an extremely simple way: either there has always been cooperation on both sides or there has been at least one defection on either side.

Appraisal of the future should satisfy two main conditions: (1) an outcome in the future should not be as valuable as that very same outcome today; and (2) the way any future outcome is viewed tomorrow should be consistent with how it is viewed today. For instance, placing some weight on tomorrow but none on the day after tomorrow cannot be consistent: indeed, since the day after tomorrow will be given some weight tomorrow and since tomorrow has some weight today, the day after tomorrow should be given some combination of these weights. The most widely accepted way to appraise future development is to discount future outcome in geometric fashion: if a player discounts tomorrow by a factor d ($0 < d < 1$) then s/he should discounts the day after tomorrow by factor d^2, the one after that by d^3, and so on…[2]

In dynamic games the sequence of future turns does not end at any preset turn, although it can end with preset probabilities. To understand the relationship between discounting and probabilistic ending, one may look at Figure 5.1. In the upper part of the figure, after the blue player chooses Play, Nature ends the game with probability $p = 0.1$ in the outcome $(10, 20)$. The game may therefore continues with probability $p = 0.9$ to reach node R. Whatever payoffs $\mathbb{E}_{\text{blue}}(\text{R})$ is expected by the blue player at node R therefore results in an expected payoff:

$$\mathbb{E}_{\text{blue}}(\text{C}) = 0.1 \times 10 + 0.9 \times \mathbb{E}_{\text{blue}}(\text{R}) \tag{3}$$

at node C from the move Play. In the lower part of Figure 5.1, the move Play yields an "instant" payoff $\text{U} = 1$ and a payoff $\mathbb{E}_{\text{blue}}(\text{R})$ discounted by factor $d = 0.9$ for a total:

$$\mathbb{E}_{\text{blue}}(\text{Play}) = 1 + 0.9 \times \mathbb{E}_{\text{blue}}(\text{R}) \tag{4}$$

[2]Some game modeling has involved "limited look ahead", meaning that a player is only concerned with a few future turns, say two for illustration. In that case, future turn number three is ignored today but will come into focus tomorrow. This is related to the idea of "bounded rationality" which advocates that decision makers can't reach the ideals of mathematical optimization.

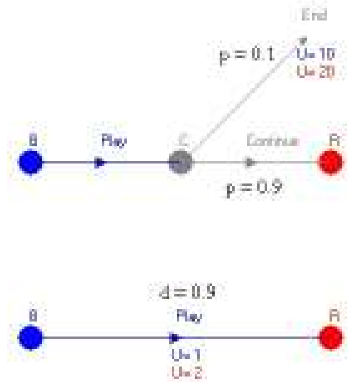

Figure 5.1: Two Equivalent Formulations of Discounting

Clearly, the results of (3) and (4) are exactly the same for any player in any such circumstances. Discounting is therefore equivalent to a chance probability of ending with an outcome that amounts to an instant payoff of taking the discounted move.

5.2 A Natural Solution Concept

In dynamic games the concept of Nash equilibrium can quickly become unsatisfactory for the same reason as those discussed in Chapter One: one can easily design suboptimal strategic plans for the future that result in an optimal plan today based on non-credible threats or pledges.

The remedy is the concept of "sequential rationality": strategies must be optimal at every single turn of the game for the player(s) deciding at that turn, given such players' *rational* expectations of the future. This avoids today's plans to rely on suboptimal, and therefore non-credible plans for tomorrow. One concept that achieves this is the SPE described in Chapter 1. That concept is still adequate for discounted repeated games, where the same static game is played over and over again. Indeed, a subgame is simply the same repeated game beginning at a later date. But when certain choices affect the very game structure that will be played tomorrow, as is often the case, it is not so simple to describe "subgames." An alternative is to use the states $\mathcal{S}_k$ defined in the partition $\boldsymbol{\mathcal{S}}$ of history described above. In that context, sequential rationality simply means that the players' choices at each state are optimal given the transition rule $\mathcal{T}$ and the expected choices at all other defined states. The resulting solution concept is called a Markov Perfect equilibrium (MPE).[3] It is easy to show that a MPE is always a SPE. It is also

[3]There is a common misconception in the literature on the meaning of MPE. The MPE was initially introduced with respect to the concept of "payoff relevant" states of the game: two states can be distinct only if the payoff structures at these two states are distinct. So, many authors draw the following *incorrect* conclusion: in a repeated game (therefore with constant payoff structure), there can be only one payoff relevant state. Therefore the only MPEs of that game are given by the repetition of a Nash equilibrium. Fudenberg and Tirole ("Game Theory", MIT Press, 1991) give a more extensive treatment of the MPE concept. They stress (pp. 513-15) that Markov strategies are based on partitions of history and that the "payoff-relevant history" is only the minimal (coarsest) *sufficient* partition. But it is in no way the *necessary* one. In a recent exposition ("A theory of regular Markov perfect equilibria in dynamic stochastic games: Genericity, stability, and purification", *Theoretical Economics*, 2010) Doraszelski and Escobar

possible to show that a MPE always exists given any partition of history $\mathcal{S}$ and corresponding transition rule $\mathcal{T}$ for a wide range of game structures.[4]

5.3 The Graph Form

The graph form can accommodate far more diverse game conditions that the simple repetition of a normal form game. In particular, it allows sequential play instead of the implicit simultaneous play of the normal form. The MPE is still the standard solution concept and one must carefully design the graph in order to represent the various possible states of memory.

5.3.1 The Dollar Auction

Professor Gotcha teaches at a state university where he thinks he is badly underpaid for his hard work. In order to supplement his income he devises the following game for his Game Theory class: he will auction a brand new $10 bill. The students will be free to bid it up, but only $1 at a time. However, there is a catch in the rules: the highest bidder will indeed get the $10 bill in exchange for his/her bid, but the second highest bidder will also pay his/her bid and will get only the professor's many thanks. When he shares his idea with a game-loving colleague, professor Gotcha adds: "At worst, it will cost me about one Dollar." Always up to the challenge, his colleague takes two Dollar bills out of his pocket and hands them to his friend saying: "Go right ahead then. If you play the game, I will give you these. So, you will now make a profit if you play the game."

Professor Gotcha teaches Game Theory using *GamePlan* and devised the model of Figure 5.2. To simplify his analysis, he assumed that two students called Blue and Red would want to play the game and that his only uncertainty is about who will move first, an issue he models by a Chance node with equal probability of either student being first to make up his/her mind about what to do. Then, he carefully distinguishes two possible turns per player. Of course, he considers the possibility of not playing the game altogether (stay) but accounts for his colleague's contribution that he can keep if he goes ahead. The discount factor $d = 0.999$ accounts for the very fast back and forth of a live auction.

---

write (p. 379): "We view a subgame perfect equilibrium of the repeated game as a Markov perfect equilibrium of a dynamic stochastic game." Despite this well published modern view of MPE the above misconception survives in many quarters.

[4]For instance, for any game on a graph that is interpretable as the repetition of a static game with discounting, the repetition of a Nash equilibrium of that static game provides a MPE.

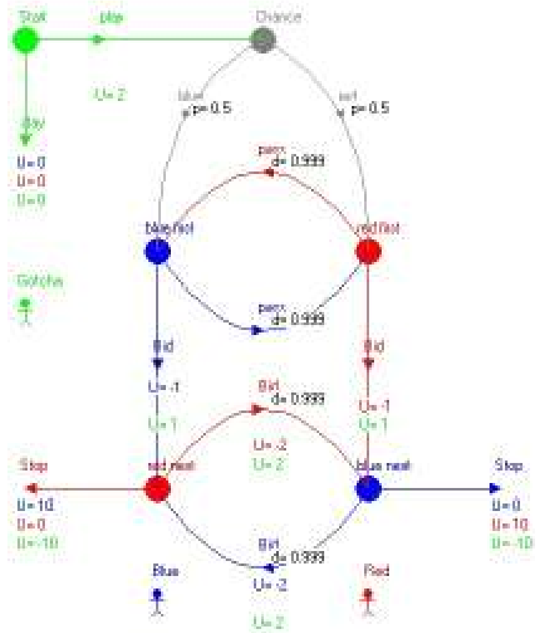

Figure 5.2: The Dollar Auction

Professor Gotcha reported a successful auction to his colleague: he walked away with $11 net (i.e., the winning bid was $11) not counting his friend's $2. In fact he had predicted that he would win at least that much with probability $p = 10.71\%$. But Dr. Gotcha had turned a blind eye to a winning strategy for his students: whoever made up his/her mind first should bid $1 and whoever would go next should abstain from bidding any further. The professor would lose $9 and be unable to brag. But he was careful not to release his lecture notes before the game.

5.3.2 Repeated Sequential Play

Sometime, repeating a game has very counterintuitive results. The alternate form of the simplest game illustrated in Figure 1.4 in Chapter One highlighted the importance of forward thinking: the threat by Red to move Left in order to deter Blue to choose Continue was dismissed as non-credible by virtue of its non-optimality. But should that simplest game be repeated, the thinking can change drastically. Figure 5.3 shows a version of that repeated game with three memory states representing the three possible plays of the one-shot game. Stop leads to the State 1 node that leads back to the Start node. But Continue followed by either of Red's moves leads to two other possible memory states: Continue followed by Left leads to State 2 in which the game unfolds again while Continue followed by Right leads to State 3. In this graph structure, the players are implicitly assumed to only keep track of the previous two moves in their interpretation of history. A richer representation of the past would require to distinguish more states. But any result obtained with only these three states would persist in a richer, but compatible definition of transitions $\mathcal{T}$.[5]

[5]By "compatible" we mean that the new partition would be a refinement of the existing partition and that the new transition rule $\mathcal{T}$ applied to the old partition would yield results consistent with the old transition rule.

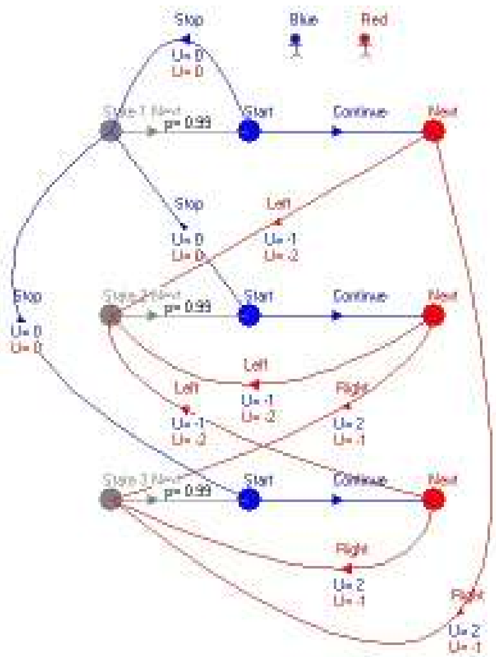

Figure 5.3: A Repeated Simplest Game

The repetition of the constituent game equilibrium $\{$Continue, Right$\}$ is a MPE of this repeated game. This is a general principle. But there are other solutions of interest. In one MPE displayed in Figure 5.4, Red will always choose Left with probability $p = 2/3$ and Right with probability $p = 1/3$.[6] As a result, and in both States 1 and 2, Blue chooses Stop with certainty. He only chooses Continue with probability $p = 50/99$ in State 3. In other words, Blue is deterred from choosing Continue unless he just observed the sequence $\{$Continue, Right$\}$. And even in that case, he still chooses Stop with probability $p = 49/99$.

The mere repetition of the static game with a minimum distinction between three possible pasts reveals a completely optimal new pattern of play. The solution of Figure 5.4 is best interpreted as the credible deterrence by Red of the move Continue by Blue, resting on her probabilistic "threat" of Left (with probability $p = \frac{2}{3}$), regardless of what happened previously. That threat is credible because it is now optimal for Red to carry it out when tested. Indeed, the threat would be tested with probability $p = \frac{50}{99}$ should state 3 be reached.

But, from a dynamic standpoint, even State 3 will not be sustained in the long run. Indeed, beginning at states 1 or 2 play will immediately lead to State 1 and will remain there. So, stepping out of State 3 immediately achieves the perpetual play of Stop. And beginning in State 3, the probability of returning to State 3 is a small $\frac{50}{297}$. The probability of staying in that state for n consecutive turns is therefore $\left(\frac{50}{297}\right)^n$ geometrically decreases with n and quickly approaches zero. At some point, Stop will be chosen in State 3 and will be maintained forever after.

[6]The reader can easily check this statement using *GamePlan*.

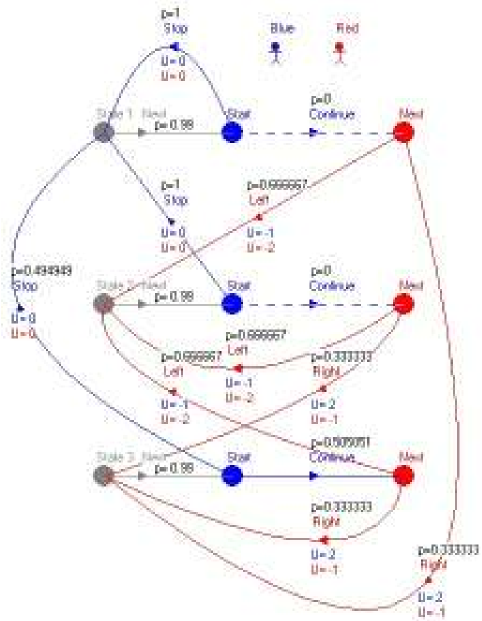

Figure 5.4: Credible Deterrence in the Simplest Game

5.4 Repeated Normal Form Games

Repeated normal form games constitute a major part of game theory. They have the double advantage of being dynamic and of relative structural simplicity. Moreover, it is hard to think of a normal form static game being played just once. In many empirically relevant game situations, the decisions will be visited several times. When the number of such decision turns has a probabilistic flavor, the repeated game framework is often appropriate.

Repeating a normal form game with finitely many memory states is particularly easy using *GamePlan*. One simply defines the constituent game and duplicates it to create as many memory states $\mathcal{S}_k$ as desired. Then, by editing the "upto" of each cell one defines the (common) discount factor as well as the transition rule $\mathcal{T}$ from that state to the next.

5.4.1 Probabilistic Return to Cooperation

The Grim Trigger obtained in the Two-states repeated Prisoner's Dilemma of Chapter 1 (Figure 1.23) is a MPE corresponding to a simplistic interpretation of history: either the two sides have always cooperated or at least one defected at least once. The scheme has the merit of providing mutual deterrence in the Cooperated state: neither side finds it appealing to defect. But there is no foregiveness built into that scheme. Once the state Defected has been reached, even by accident or misunderstanding, it is maintained forever according to the equilibrium.

So, what if some device could re-establish the mutual trust implicit in the Cooperated state? This could be a good deed outside the game, a benevolent third party, or even some "sign from the gods." We will model this possibility as a random move by Chance in Figure 5.5.

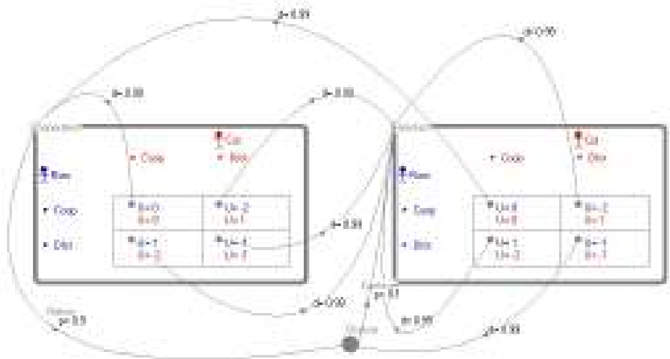

Figure 5.5: Chance Return to Cooperation

A MPE akin to the Grim Trigger emerges from this slight modification. The difference is that the Dfct-Dfct cell in the Defected state now leads to the chance move that re-establishes cooperation with probability $p = \frac{9}{10}$. So, cooperation will be sustained in the Cooperated state and will be re-established quite soon from the Defected state. From a dynamic standpoint, this is a much more promising plan than the Grim Trigger. However, it leaves open the question of how such a device can be engineered and by whom.

5.4.2 Four-States Prisoner's Dilemma

The re-establishment of cooperation by the players' rational play after some episode of retaliations would be far more persuasive than the outside mechanism described in the previous section. Put simply, the question is: can one devise a completely rational plan that achieves deterrence but will eventually reestablish cooperation after any episode, intentional or accidental, of defection in the repeated Prisoner's Dilemma?

Trusting in ancient wisdom, one may consider the simple strategy "an eye for an eye.." as a candidate for achieving just that. The strategy is usually called "Tit-for-tat" (TFT) in the Game Theory literature. Formalizing it with GamePlan first requires distinguishing enough states of memory to represent all possible histories and all reactions according to TFT played against TFT. This is achieved by partitioning all histories into four states: (1) CC: all histories ending in bilateral cooperation; (2) DD: all histories ending in bilateral defection; (3) DC: all histories ending with defection by Row and cooperation by Column; and (4) CD: all histories ending in coopearation by Row and defection by Column. The transitions from state to state are then obvious. The resulting *GamePlan* model is shown in Figure 5.6.

All plays of $\{$Coop,Coop$\}$ lead to state CC, all plays of $\{$Dfct,Dfct$\}$ lead to state DD, and unilateral defection leads to one of the other two states as appropriate. A uniform discount factor $d = 0.99$ is applied.

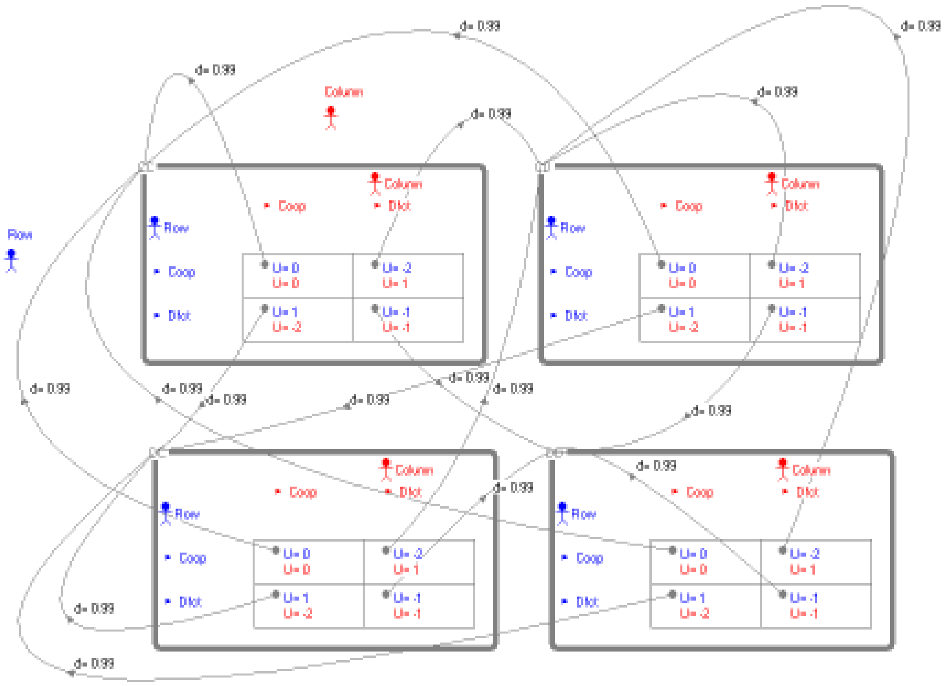

Figure 5.6: A 4-States Repeated Prisoner's Dilemma

Solving with *GamePlan* is somewhat disappointing: the standard repetition of the Nash equilibrium {Dfct,Dfct} of the static game naturally emerges as a MPE as well as our familiar Grim Trigger. But nowhere does one find TFT in the list. The closest is:

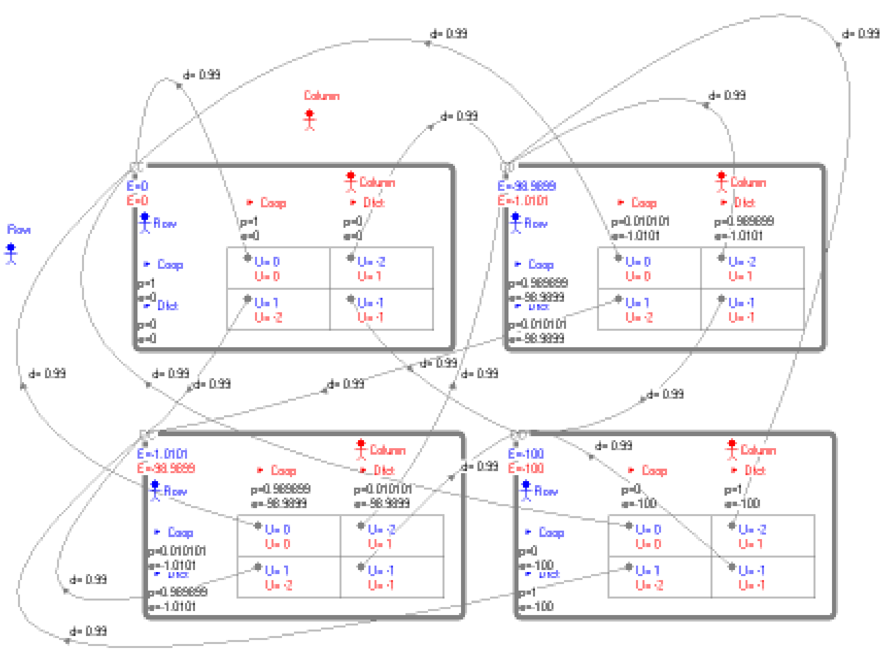

Figure 5.7: A Tit-for-tat Like Solution

Indeed, the TFT pattern appears in CC and DD, but only with probabilities in the two Unilateral Dfct states but not in the No Unilateral Dfct one where the previous defector cooperates with $p \simeq 0.99$ probability and the agrieved previous cooperator defects with the same probability $p \simeq 0.99$. In order to understand exactly why TFT vs TFT doeas not form a MPE, a formal analysis is called for.

If TFT is assumed to be played against itself, any instance of {Coop,Dfct} at one turn will be followed by {Dfct,Coop} at the next turn, followed by {Coop,Dfct} again, and so on. So, it is easy to obtain the expected payoffs for Row at states CC, CD and DC:

$$\mathbb{E}_{\text{Blue}}[\text{CC}] = 0 + \mathbb{E}_{\text{Blue}}[\text{CD}] = \tag{5a}$$

$$\mathbb{E}_{\text{Blue}}[\text{CD}] = 1 + d \times \mathbb{E}_{\text{Blue}}[\text{DC}] = 1 + d(-2 + d \times \mathbb{E}_{\text{Blue}}[\text{CD}]) = \tfrac{1-2d}{1-d^2} \tag{5b}$$

$$\mathbb{E}_{\text{Blue}}[\text{DC}] = -2 + d \times \mathbb{E}_{\text{Blue}}[\text{CD}] = \tfrac{d-2}{1-d^2} \tag{5c}$$

For TFT to be a best reply to TFT at CC it is necessary that choosing Dfct at CC or Coop at CD is counterproductive for Blue. Formally:

$$\mathbb{E}_{\text{Blue}}[\text{Dfct}|\text{CC}] = 1 + d \times \mathbb{E}_{\text{Blue}}[\text{DC}] = \tfrac{1-2d}{1-d^2} \leq 0 = \mathbb{E}_{\text{Blue}}[\text{CC}] \tag{5d}$$

and $$\mathbb{E}_{\text{Blue}}[\text{Coop}|\text{CD}] = 0 + \mathbb{E}_{\text{Blue}}[\text{CC}] = 0 \leq \tfrac{1-2d}{1-d^2} = \mathbb{E}_{\text{Blue}}[\text{CD}] \tag{5e}$$

Clearly (5d) and (5e) can only hold for the very unlikely value $d = \frac{1}{2}$. So, in general, TFT vs TFT cannot form a MPE.

In general, it is reasonable to assume that d is relatively high so that (5d) holds. The resulting failure of (5e) can be understood in the following light: Suppose that Column unilaterally defected. Then, at the very moment when Row is preparing to retaliate, Column comes to her and makes the following plea: "look, this was all a big mistake. I did not intend to defect on you. I just did it by accident. Please forgive me and skip the retaliation. In fact, if you don't skip it, look at what will happen: I will play Coop according to TFT since you just cooperated. So, your retaliation will simply exchange our roles in the above calculations and you will face an expected payoff:

$$\mathbb{E}_{\text{Blue}}[\text{Dfct}|\text{CD}] = \tfrac{1-2d}{1-d^2} < 0 = \mathbb{E}_{\text{Blue}}[\text{Coop}|\text{CD}] \tag{6}$$

So, concludes Column, you are better off forgiving my "mistake". If Row follows the advice, Column will undoubtedly make such further "mistakes" so that TFT will lose any deterrent credibility against the "mistaken" TFT.

$$\mathbb{E}_{\text{Red}}[\{\text{Coop,Dfct}\}|\text{TFT}] = \tfrac{1-2d}{1-d^2} < 0 = \mathbb{E}_{\text{Red}}[\{\text{Coop,Coop}\}|\text{TFT}]$$

Ancient wisdom does not always work in Game Theory, or does it? There is, indeed, a remedy that requires a subtle twist of interpretation discussed in homework 5.7.9. It involves a form of moral judgement.

5.4.3 An Environmental Treaty

The neighboring states of Megasmog and Pristina have a serious dispute that threatens their longstanding peace: the Blue River that flows from the Megasmog industrial region in the North, along their common border to the South, has become increasingly polluted. Fortunately for Megasmog, it has access to the sources of the river and therefore enjoys a clean water supply. But Pristina's citizens are reduced to filter their water of to buy bottled water produced by the Megasmog Upper River Water Company.

Some of Pristina's businesses are pushing to relax its strict anti-pollution laws in order to retaliate. But polluting the environment further would be to the detriment of both sides. Game Theory Associates (GTA), a consulting firm, describes the situation by the normal form game of Figure 5.8.

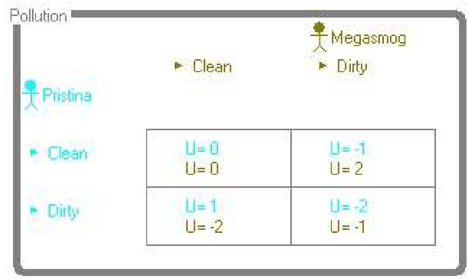

Figure 5.8: The Pollution Game

The situation appears hopelessly disadvantageous for Pristina. But GTA contends that an environmental treaty that would maintain clean policies on both sides is entirely possible. It is only a matter of design. After long negotiations, the two sides agree to consider three "states" of the treaty: Compliance, Megasmog non-compliance, and Pristina non-compliance. The non-compliance state will be reached by the side that is found to unilaterally dirty the environment. The two sides will then remain in that state for a few turns before returning to Compliance. While in non-compliance, the state responsible will clean the environment while the other will be expected to play Dirty. Return to Compliance will be decided by an independent panel with a given probability p. GTA has proposed the model of Figure 5.9.

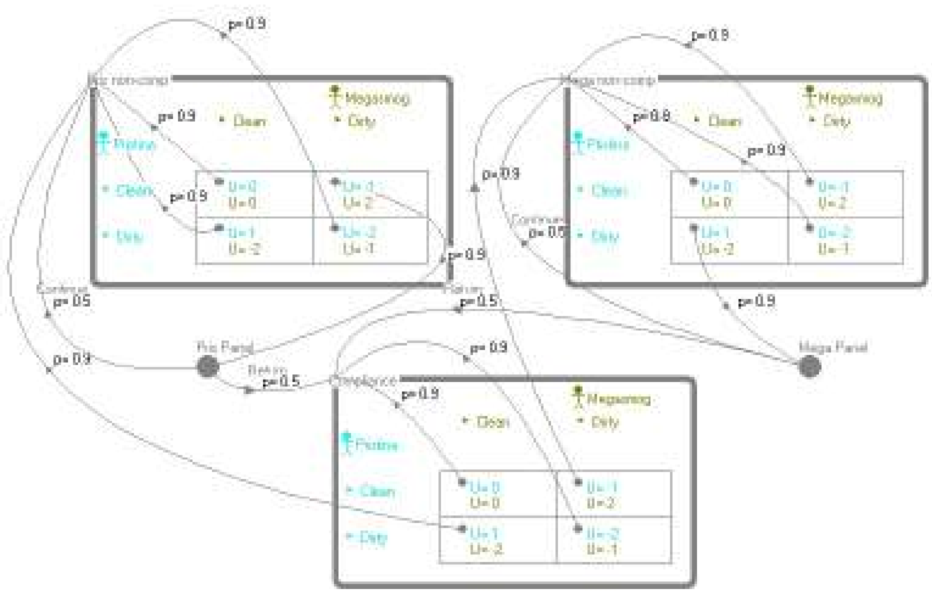

Figure 5.9: An Environmental Treaty

The Megasmog and Pristina delegations to the talks find that proposition dubious, to say the least. They immediately assail the GTA representative, Dr. Green, with questions: why should Pristina dirty the environment when Megasmog is in non-compliance since their objective is to protect the environment? Why should the panels decide on a probabilistic return to compliance rather than after a fixed period of time? Dr. Green explains that this doesn't make any difference. Even if Pristina is not required to dirty the environment while Megasmog is in non-compliance, it will still do so under the pressure of its business community, since that is allowed by the treaty terms. And if it's not allowed, the treaty has no teeths. And as for a fixed number of turns, it makes no difference since a probability of return to compliance defines an expected number of turns of non-compliance.

Dr. Green explains that this setup yields a Markov perfect equilibrium (MPE) with Compliance as a stable steady state. You can adjust some of the parameters, she adds, such as the probability of return, but the treaty should succeed.

5.4.4 The Tragedy of the Commons

There are several generalizations of the two-player prisoner's dilemma to three or more players even when symmetry is preserved. It all depends on the effect of accumulating defections. In the game of Figure 3.6, in Chapter Three, a unanimous defection brings the worst possible result for all three players. When played just once, this game has three very symmetrical pure Nash equilibria: one side cooperates while the other two defect. But any of the three sides can be the "victim" and it is therefore hard to predict whom that will be when the game is played for real. Such a situation has been described as the "Tragedy of the Commons", a social dilemma involving a population of self-interested decision makers whose rational individualistic behavior can lead to social catastrophe.

The repetition of that game can easily yield cooperation, depending on how the memory states and the state transitions are defined. For instance, one can create four memory states: Cooperation and one for each possible victim. When all sides cooperate or all simultaneously defect, the state of Cooperation endures. Any deviation by one or two sides lead to a victim state where the victim is expected to defect in retaliation while at least one of the defectors will cooperate. With high enough discount factors this yields a MPE where full cooperation endures. However, the transitions can be engineered in such a way that cooperation will be quickly reestablished rationally (see homework..).

5.4.6 Payoff Relevant States

The original definition of the Markov Perfect Equilibrium referred to the concept of "payoff-relevant" states. The idea was that if two distinct histories give rise to distinct move and/or payoff structures then they could not be merged into a same state. A simple example that illustrates this concept is inspired by the Egyptian Dilemma mentionned in the homework section of Chapter One. There are two states of the game: Democracy and Autocracy. In the first state, the democratically elected government is dominated by a Party with a radical religious base. So it can implement a radical or a liberal policy. The Army can either submit to the elected government or take over by making a coup. In Autocracy, the Party can either submit to the Army or rebel and the Army can tolerate or repress the Party. The payoff structure is inherently different in the two states. One possible definition is in Figure 5.10.

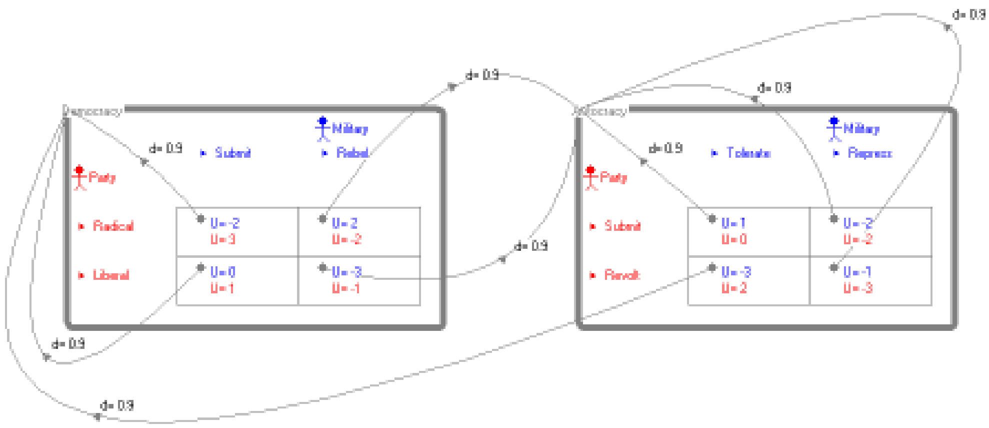

Figure 5.10

The transitions reflect a possible evolution according to the two sides' choices. There are others (see homework).

5.4.7 Folk Theorems

In mathematics, folk theorems are well known results whose authorship is unclear. In game theory the term refers to various statements about simple equilibria of repeated games. Perhaps the simplest and most powerful Folk Theorem concerns the Grim Trigger: suppose that a constituent game admits a strategy profile that is not in equilibrium but yields a strictly better outcome than a Nash equilibrium of the same game, for all players. Then, if the players have enough concern for the future (i.e. high enough discount factors), the Grim Trigger that sustains the better outcome through the threat of perpetual reversion to the Nash equilibrium forms a MPE in the repeated game. The typical example is given in Figure 1.24 in Chapter One, but there are numerous other cases.

5.5 Repeated Continuous Games

Continuous games such as the duopoly and oligopoly studied in Chapter One and Chapter Three are great candidates for repetition. Indeed, many economic games are by nature repeated in time and should therefore be studied in that perspective. The Cournot duopoly and oligopoly give rise to equilibria that are not as efficient as what could be achieved with some collusion (see homework.) So, they lend themselves to the same improvements using trigger schemes. However, trigger schemes require a clear understanding between the players of what point within an entire continuum will be chosen as the target to maintain as well the exact schedule of retaliation needed to sustain it. This creates some credibility issues as far as applications are concerned.

So, the following question arises: is it possible to design perfect equilibria that sustain an efficient outcome, maintain that outcome dynamically and do not entail anything more that unilateral pledges or threats to formulate? The answer is yes, but it requires the technical developments of the next section.

5.5.1 The Decomposition Theorem

Let $x_i^t \in A_i$ denote Player i's decision in her action space A_i and let $X_{-i}^t \in \underset{j \neq i}{\times} A_j$ denote all other players' decisions in their own action spaces, at turn t. Further let $U_i(x_i^t, X_{-i}^t)$ denote i's constituent game payoff resulting from such decisions at turn t. Player i' objective in the discounted repeated game (with discount factor δ_i for i) is to maximize, at each turn t, the discounted sum of present and future payoffs:

$$E_i(\xi_i^t, \Xi_{-i}^t) = \sum_{s=0}^{\infty} \delta_i^s U_i(x_i^{t+s}, X_{-i}^{t+s}) \tag{4}$$

where $\xi_i^t = \{x_i^{t+s}\}_{s \in \mathbb{N}}$ denotes Player i's present and future choices and $\Xi_{-i}^t = \{X_{-i}^{t+s}\}_{s \in \mathbb{N}}$ denotes all other players' expected present and future choices. The history of the game evolves according to $h^{t+1} = h^t \cup (x_i^t, X_{-i}^t)$.[7] More generally, if the game has memory states and a transition rule $\mathcal{T}$, one has:[8]

$$h^{t+1} = \mathcal{T}(h^t, (x_i^t, X_{-i}^t)) \tag{5}$$

If $h^t \in \boldsymbol{\mathcal{H}}$ denotes the history, or state, of the repeated game at turn t, a strategy for Player i is a map

$$\psi_i : h^t \in \boldsymbol{\mathcal{H}} \to x_i^t \in A_i$$

We also denote by $\Psi = (\psi_i, \Psi_{-i})$ a strategy profile such that $x_i^t = \psi_i(h^t)$, and similarly for all players. One has:[9]

Theorem 5.1: Ψ is a MPE if and only if there exists for each player i two maps $\pi_i \geq 0$ and g_i (of any sign) satisfying

$$g_i(X_{-i}^t, h_t) - \pi_i(\xi_i^t, X_{-i}^t, h_t) = U_i(x_i^t, X_{-i}^t) + \delta_i g_i(\Psi_{-i}(h^{t+1}), h_{t+1}) \tag{6a}$$

with $\quad \pi_i(\xi_i^t, \Psi_{-i}(h_t), h_t) = 0 \qquad$ if $x_i^{t+s} = \psi_i(h^{t+s})$ for all $s \geq 0$ (6b)

This is merely a version of the Bellman Equation of Dynamic Programming with a twist that will be extremely useful in applications. But each of the two funtions involved in (6a) finds an interesting interpretation: suppose that one chooses $\pi_i \equiv 0$. Then, whatever Player i could gain by a choice $x_i^t \neq \psi_i(h^t)$ in $U_i(x_i^t, X_{-i}^t)$ would be cancelled *exactly* by the term $\delta_i g_i(\Psi_{-i}(h^{t+1}), h_{t+1})$ according to the other players reaction $\Psi_{-i}(h^{t+1})$, since the two terms must add up to $g_i(X_{-i}^t, h_t)$ which is *independent* of x_i. For that reason, the g_i term has been called the "countervailing" part of Ψ_{-i}. Equilibria with that property are entirely feasible and have been called countervailing. As shown in examples below, they can arise spontaneously from threats and pledges and need not require any coordination in the strategic choice of the players. When it is not identically nil, the π_i term has the effect of holding the players to a specific strategic choice ψ_i as suggested by (6b). This has been called the "coercive" part of Ψ_{-i}. This may be a desirable feature, for instance if the equilibrium is the result of a treaty design.

In practice, the theorem is applied with history $\boldsymbol{\mathcal{H}}$ partitioned into a few relevant states. In some interesting cases, there can even be a single trivial "null" state together

[7]If the game begins at time $t = 0$ one has $h_0 = \emptyset$.
[8]$\boldsymbol{\mathcal{T}}(h_t, \Psi(h_t)) = h_t \cup \Psi(h_t)$ can be viewed as a trivial case of transition rule in $\boldsymbol{\mathcal{H}}$.
[9]This theorem first appeared in Langlois & Langlois (1996).

with the countervailing condition $\pi_i \equiv 0$ so that the formula (6a) reduces to the very simple condition:

$$g_i(X_{-i}^t) = U_i(x_i^t, X_{-i}^t) + \delta_i g_i(\Psi_{-i}(X^t)) \tag{7}$$

In that case, g_i can easily be determined by reference to unilateral threats or pledges and (7) can be explicitly solved for Ψ. The next two sections give some examples.

5.5.2 Unilateral threats and pledges

The normal form Prisoner's Dilemma of Figure 1.23 in Chapter One can be reinterpreted as a continuous game with choices $x_i \in [0, 1]$ and payoffs

$$U_i(x_i, x_j) = x_i - 2x_j \tag{8}$$

where x_i is interpretable as a "level of defection." The four corners of the resulting (square) action space provide exactly the same payoffs as in Figure 1.23. This continuous Prisoner's Dilemma can serve as a generalization of the discrete version. Its repetition with discount factor δ (common to the two sides for simplicity) is a typical case where Theorem 5.1 can be applied. It is easiest to construct countervailing equilibria by setting $\pi_i \equiv 0$. It is usually quite easy to then modify that equilibrium into a coercive one if need be. In the continuous case is also quite helpful to define the state of the game as null, meaning that reactions are limited to the last play of the constituent game. In this two player case, equation (7) then reduces to:

$$g_i(x_j^t) = x_i^t - 2x_j^t + \delta g_i(\psi_j(x_i^t, x_j^t)) \tag{9}$$

So, if one "knows" g_i and if it is monotonic, it is easy to reconstruct ψ_j by simply solving (9). The question is how g_i can be known? It turns out that g_i can be completely determined by unilateral threats or pledges made by player j. For instance, suppose that Player j offers to progressively reciprocate i's full cooperation by cutting her defection level in half at each turn. She is pledging:

$$\psi_j(x_i^t = 0, x_j^t) = x_j^t \div 2 \tag{10}$$

But this entirely determines g_i in (9). Indeed, one can write:

$$g_i(x_j^t) = \; -2x_j^t + \delta g_i(x_j^t \div 2) = \tfrac{-4}{2-\delta} x_j^t \tag{11}$$

Replacing in (9) yields the formula:

$$\psi_j(x_i^t, x_j^t) = \tfrac{(2-\delta)x_i^t + 2\delta x_j^t}{4\delta} \tag{12}$$

Provided that $\delta \geq \frac{2}{3}$, the map defined by (12) is indeed a strategy since it takes all its values in $[0, 1]$.[10] Player i can formulate independently his own pledge or threat and obtain the corresponding strategy in similar fashion. If this also yields a true strategy as in (12), the result is a MPE. And should Player i make the symmetric pledge, he would obtain the symmetric strategy and the two sides would find themselves in an MPE with an elegant property: it sustains and dynamically re-establishes cooperation after any episode of unilateral or bilateral defection.

[10] Its maximum value is $\psi_j(1, 1) = \frac{2+\delta}{4\delta} \leq 1$ provided $\delta \geq \frac{2}{3}$.

Instead of making the above pledge of partial reciprocation, Player j may instead make a threat of partial retaliation. For instance, she can threaten progressive retaliation to full defection by cutting in two her current distance to full defection. This means:

$$\psi_j(x_i^t = 1, x_j^t) = 1 - (1 - x_j^t) \div 2 = (1 + x_j^t) \div 2 \qquad (13)$$

Again, this determines g_i in the countervailing case:

$$g_i(x_j^t) = x_i - 2x_j^t + \delta g_i((1 + x_j^t) \div 2) = \frac{2-3\delta}{(1-\delta)(2-\delta)} - \frac{4}{2-\delta} x_j^t \qquad (14)$$

Replacing in (9) yields:

$$\psi_j(x_i^t, x_j^t) = \frac{(3\delta-2)+(2-\delta)x_i^t+2\delta x_j^t}{4\delta} \in [0, 1] \qquad (15)$$

which is a strategy as long as $\delta \geq \frac{2}{3}$. Again, Player i can formulate his own pledge or threat independently. Regardless of what threat he adopts, as long as it also defines a true strategy, one again obtains an MPE. But if i adopts the symmetric threat, one obtains a not very attractive MPE: it sustains and dynamically re-establishes full defection after any episode of unilateral or bilateral cooperation.[11]

This is not to say that threats are inappropriate for all sorts of games. It only illustrates how unilateral statements coupled with a countervailing assumption can yield MPEs with extremely different dynamic properties.

5.5.3 Reaction Function Equilibria

In the late 1960's, it was conjectured that Cournot's reaction functions (see Section 1.2.6) could be replaced by a MPE that would promote a more cooperative outcome than the Nash-Cournot equilibrium.[12] The conjecture was proven correct in the early 1990's.[13] The technique is illustrated with the simplest revenue model of Chapter One.

If the two sides of the duopoly were to collude and "fix" prices, they could agree to produce equal levels $q_1 = q_2 = q$ and split the proceeds. They would thus jointly maximize

$$qf(2q) = q(60 - 2q)$$

by choosing $q = 15$ (rather than the Nash-Cournot equilibrium $q = 20$.) They would each enjoy the collusive revenue $qf(2q) = 15 \times 30 = 450$ (instead of $20 \times 20 = 400$.) In order to obtain a MPE that would achieve a collusive outcome, one can define a countervailing g_i function and solve for the reaction function ψ_j

$$g_i(q_j) = q_i(60 - q_i - q_j) + \delta g_i(\psi_j) \qquad (16)$$

A simple choice is

$$g_i(q_j) = \lambda - \mu q_j$$

---

[11] One easily obtains $x_i = x_j = 1$ as the only steady state. To show its dynamic stability, one obtains the Jacobian matric DΨ and find its eigenvalues $\lambda = \frac{1}{2} \pm \frac{2-\delta}{4\delta} \in (0, 1)$, provided $\delta > \frac{2}{3}$. By a standard theorem on discrete dynamical systems, the iteration of the map Ψ converges to the steady state.

[12] That conjecture was expressed by James W. Friedman.

[13] Langlois & Sachs (1993) and Friedman & Samuelson () independently achieved the result.

which yields by (16)

$$\lambda - \mu q_j = q_i(60 - q_i - q_j) + \delta(\lambda - \mu\psi_j) \tag{17}$$

or $$\psi_j(q_i, q_j) = \frac{q_i(60-q_i-q_j)+\mu q_j-(1-\delta)\lambda}{\delta\mu} \tag{18}$$

One can now impose on (17) the condition that some $q^* = \psi_j(q^*, q^*)$ is a steady state of the MPE. This yields a relation between λ and μ. It is entirely possible to choose these in such a way that the collusive outcome $q^* = 15$ be the desired steady state. Unfortunately, this fails to provide dynamic stability to that steady state, a very desirable property. Instead, one can choose an intermediate value such as $q^* = 16$. Setting $\mu = q^* = 16$ then yields (with, say $\delta = 0.9$) $\lambda = 4,736$ and

$$\psi_j(q_i, q_j) = \frac{q_i(60-q_i-q_j)+\mu q_j-(1-\delta)\lambda}{\delta\mu} = \frac{5}{72}\Big(q_i(60 - q_i - q_j) + 16q_j - 473.6\Big) \tag{19}$$

One can verify that ψ_j, together with its symmetric ψ_i have $(q_i, q_j) = (q^*, q^*)$ as a dynamically stable steady state.[14] As a result, they map a neighborhood Ω of that point into itself. The decomposition theorem then yields a MPE based on the distinction between two states: whenever $(q_i, q_j) \in \Omega$ the state is $\mathcal{S}_0$. Otherwise the state is $\mathcal{S}_1$. In state $\mathcal{S}_0$ play is expected to be according to (19) with the corresponding $g_i(q_j, \mathcal{S}_0)$. Otherwise, $\psi_i = \psi_j = 20$ with the corresponding $g_i(q_j, \mathcal{S}_0)$ obtained by the expectation of $q = 20$ forever.[15]

5.6 Attrition and Bargaining

An interesting twist on repeated games arises when the players have the capability to end the game by their own choice. The two major examples are wars of attrition and the repeated game model of bargaining.

5.6.1 The War of Attrition

Consider a two-player repeated game where a player's choice at their turn is whether to end the game in a loss for themselves, and a gain for the other, or continue playing the game at a cost, thereby giving the other side the symmetric choice at the next turn. The situation is best pictured as the most basic graph form of Figure 5.11.

---

[14] The eigenvalues of the Jacobian matrix $D[\psi_i, \psi_j]$ at $(16, 16)$ are $\lambda_i = \lambda_j = \frac{5}{6}$, less than one in absolute value. This implies the dynamic stability of the steady state under the discrete dynamics defined by Ψ.

[15] By involving a coercive term π_i, as in (6b), it is possible to construct reaction function equilibria that support the collusive point $(15, 15)$.

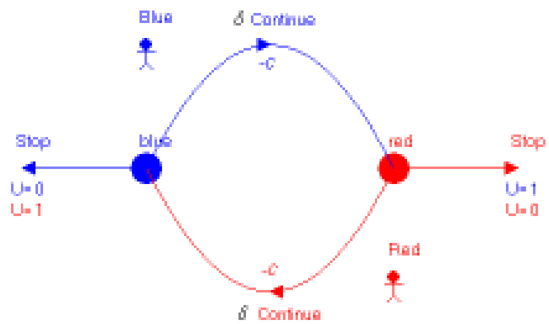

Figure 5.11: The War of Attrition

This game has three MPEs: in two pure equilibria, one side chooses Stop while the other chooses Continue. In a more interesting symmetric MPE, each side continues with a probability that is linked to the game parameters. Here, it is assumed that some prize of value $U = 1$ is at stakes and that the player who chooses Stop hands it to the other (and keeps nothing.) It is easy to solve for the mixed MPE: by symmetry, we may assume a common probability p of continue (and $(1-p)$ of Stop.) At node blue, for instance, player Blue anticipates an expected payoff for Continue:

$$\mathbb{E}_{\text{Blue}}(\text{Continue}|@\text{blue}) = -c + \delta\mathbb{E}_{\text{Blue}}(@\text{red})$$

$$= -c + \delta\big(p\delta\mathbb{E}_{\text{Blue}}(@\text{blue}) + (1-p) \times 1\big) = \frac{\delta(1-p)-c}{1-\delta p}$$

In order for the probability p to be rational for Blue at node blue, one must have $\mathbb{E}_{\text{Blue}}(\text{Continue}|@\text{blue}) = \mathbb{E}_{\text{Blue}}(\text{Stop}|@\text{blue}) = 0$, or:

$$p = 1 - \frac{c}{\delta} \tag{20}$$

So, as long as $c < \delta$, the War of Attrition can continue rationally on with that probability at each turn.

5.6.2 The Rubinstein Bargaining Model

The bargaining problem is as old as Game Theory. In fact, John Nash made his contribution with what is now known as the "Nash Bargaining Solution." This was an axiom-based formula for what bargain should emerge given certain parameters. What became known as the "Nash Program" is the goal of explaining such outcomes through the non-cooperative game theoretic approach. Rubinstein's bargaining model is typical of the Nash Program. The game structure is in fact the same as the above War of Attrition one. The only difference is that the payoff to each side is the result of an offer by the other at the previous turn. The result is shown in Figure 5.12:

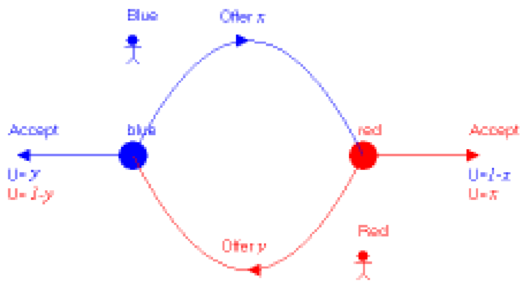

Figure 5.12: The Rubinstein Bargaining Model

Again, the future is discounted by factor δ and player Blue can make the following calculation at node blue: $\mathbb{E}_{\text{Blue}}(\text{Offer } x) = \delta\mathbb{E}_{\text{Blue}}(\text{@red})$. But now, one focuses on adjusting x and y to reach the earliest possible bargain since the others will be less valuable since discounted further. This implies an immediate acceptance by each side of the corresponding current offer.[16] This yields:

$$\mathbb{E}_{\text{Blue}}(\text{Accept } y) = y = \mathbb{E}_{\text{Blue}}(\text{Offer } x) = \delta(1 - x)$$

With the symmetric condition $x = \delta(1 - y)$, one finds the optimal bargain $x = y = \frac{1}{1+\delta}$.

5.7 Homework

5.7.1 The Repeated Nash Equilibrium as MPE

Argue in your own word why the repetition of a same Nash equilibrium of a repeated constituent game is a MPE of the discouted repeated game. Hint: consider an arbitrary number of memory states and pick any one of them. If the play of the Nash equilibrium is expected in all other memory states, what is best in the picked memory state?

5.7.2 The Best Environmental Treaty

In probability theory, if an event recurs with fixed probability p and ends with probability $(1 - p)$, the expected number ν of event turns is defined by:

$$\nu = (1 - p)\sum_{n=1}^{\infty} np^{n-1} = \frac{1}{1-p}$$

(a) What probability p corresponds to an expected number of 10 turns?

(b) What is the maximum probability p that is compatible with the success of the environmental agreement described in §5.4.3?

5.7.3 The Repeated Battle of the Sexes

Construct a repeated game model of the Battle of the Sexes and obtain a MPE that will sustain the alternate and joint choice of Ballet and Fight by the two players.

5.7.4 Guilt in the Prisoner's Dilemma

---

[16] Rubinstein's argues that any best SPE outcome at any future point in time will be the same, only discounted and therefore less valuable. The best is therefore what is immediately acceptable by both sides.

It was established in the text that the Tit-for-tat strategy (played against itself) does not provide a credible threat since it is not optimal to carry it out when tested. It was also established that credible threats do exit although they either grim or involve debatable outside mechanisms. Could there be more sophisticated designs that would not just deter defection credibly but promote it and even re-establish cooperation after any episode of defection? The answer is yes and, although there are many such schemes possible, there is a particular instructive one called "Contrite Tit-for-Tat": The idea is to introduce a concept of guilt in the players' representation of history.

There are three possible states in the game: one side or the other is guilty (of inappropriate defection) or neither side is. One becomes guilty by defecting unilaterally on a non-guilty side. One remains non-guilty when defecting on a guilty side in retaliation for its bad deed. And one always becomes non-guilty by cooperating.

You will edit the GamePlan model of Figure 5.7 to represent this interpretation of history. You will then solve the game for pure strategy equilibria and comment.

5.7.5 The Repeated 3-Way Prisoner's Dilemma

Construct a repeated game model of the 3-player Prisoner's dilemma with 4 states defined as follows: DCC for "only Blue defected," CDC for "only Red defected," CCD for "only Green" defected, and Neither for "neither of the other three states." Define transitions between states accordingly and solve for pure equilibria. Comment on your results.

5.7.6 The Egyptian Dilemma

First solve the model of Figure 5.10 in section 5.4.6. Then modify it by introducing a chance node (called Revolution) that a revolution will succeed or fail after the pair of choices $\{$Revolt,Repress$\}$. Success would return to Democracy while failure would return to Autocracy. Vary their respective probabilities and comment on your results (no formal analysis is required).

5.7.7 Collusion in Oligopoly

Generalize the construction of section 5.5.3 to the case of three oligopolists (i, j, k).

(a) Show that the Nash-Cournot equilibrium is at $q_i = q_j = q_k = 15$.

(b) Show that the collusive point is at $q_i = q_j = q_k = 10$.

(c) Let $g_i(q_j, q_k) = \lambda - \mu(q_j + q_k)$, and symmetrically for j and k. Using $\mu = 11$ and $\lambda = 3,212$, solve a system of three equations in three unknowns (ψ_i, ψ_j, ψ_k) (here written for i):

$$g_i(q_j, q_k) = q_i(60 - q_i - q_j - q_k) + \delta g_i(\psi_j, \psi_k)$$

Verify that $q_i = q_j = q_k = 11$ is a steady state for the reaction functions (ψ_i, ψ_j, ψ_k).[17]

[17]It is possible to show that this steady state is dynamically stable under the dynamics defined by the reaction functions.

5.7.8 The Cuban Missile Crisis

Nuclear crises were described by Herman Kahn as a Game of Chicken. Consider the following continuous game utility functions (and symmetrically, by exchanging i and j):[18]

$$U_i(x_i, x_j) = x_i - x_j - 2x_i x_j$$

This is the continuous extension of the Chicken (normal form) game of Figure 5.13 where $x_i = 0$ means Swerve and $x_i = 1$ means Drive On. In the context of a nuclear crisis, one may interpret x_i as a "level of escalation."

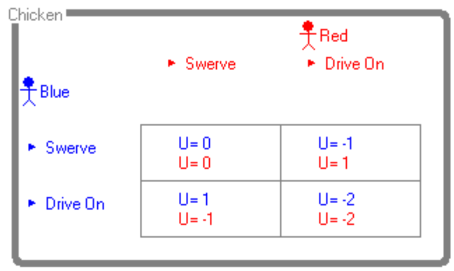

Figure 5.13: The Game of Chicken

Suppose that the future is discounted by $\delta > \frac{1}{2}$. Further assume that the US (as j) offers to reciprocate full cooperation ($x_i = 0$) by the Soviet Union (SU) according to the formula:

$$\psi_j(x_i = 0, x_j) = x_j \div 2\delta$$

In essence, if δ is close enough to 1, this means that US will cut its level of escalation in half at each turn, should SU stick to $x_i = 0$. It is a pledge of incremental reciprocation.

(a) Show that this pledge is equivalent to a countervailing strategy with $g_i(x_j) = -2x_j$. Hint: assume $g_i(x_j) = -\mu x_j$, solve for ψ_j and apply the above condition.

(b) Argue that the best SU can hope for in the long run is to maintain full cooperation. Hint: show that, whatever steady state (x_i, x_j) is ever reached, it will have to satisfy

$$(2\delta - 1)x_j = (1 - 2x_j)x_i$$

So, the best "long term" $g_i(x_j) = -2x_j$ can only occur if $x_i = 0$.

(c) Show that the very same conclusions can be reached if US instead threatens incremental retaliations $\psi_j(x_i = \frac{1}{2}, x_j) = (1 + x_j) \div 4\delta$.

[18]One justification for such a structure was offered in Langlois (1991.)