Advanced microeconomics 3: Game Theory

Slide set 2: Dynamic games

Pauli Murto

Aalto University

February 3, 2023

Outline

- We now move to dynamic games
- We focus especially on Nash equilibrium refinements induced by sequential rationality: sub-game perfect equilibrium and sequential equilibrium
- This slide set covers weeks 3-4 of the course
- Material: MWG Chapter 9, Mailath Ch. 1.3, 2.2-2.3, 5
- Other relevant sources: Fudenberg-Tirole Ch. 3-4, 8.3, Osborne-Rubinstein Ch. 6-7, 12, Myerson Ch. 4

Example: predation

- An entrant considers entry into an industry with a current incumbent firm
- Entry costs 1 unit
- Monopoly profit in the industry is 4
- If entry takes place, the monopolist can either accomodate or fight
- Accommodation splits monopoly profits, whereas fighting gives zero profit to both firms
- Will entrant enter, and if so, will incumbent fight or accomodate?

- Normal form representation of the game:

	Fight if entry	Accommodate if entry
Enter	$-1, 0$	$1, 2$
Stay out	$0, 4$	$0, 4$

- There are two Nash equilibria: (Enter, Accommodate) and (Stay out, Fight if entry)
- Is one of the two equilibria more plausible?

Example: quality game

- A producer can produce an indivisible good, and choose either high or low quality
- Producing high quality costs 1 and bad quality 0
- Buyers values high quality at 3 and bad quality at 1
- For simplicity, suppose that good must be sold at fixed price 2
- Which quality will be produced and will the buyer buy?

- Normal form representation of the game:

	High quality	Low quality
Buy	$1, 1$	$-1, 2$
Do not buy	$0, -1$	$0, 0$

- Only one Nash equilibrium (Do not buy, Low)
- What if seller moves first?
- What if buyer moves first?
- What if seller moves first, but quality is unobservable?

Example: Stackelberg vs. Cournot

- Consider the quantity setting duopoly with profit functions

$$\pi_i\left(q_i, q_j\right) = q_i(1 - q_1 - q_2), \text{ for all } i = 1, 2.$$

- Suppose the players set their quantities simultaneously (Cournot model). The unique Nash equilibrium is:

$$(q_1^*, q_2^*) = \left(\frac{1}{3}, \frac{1}{3}\right),$$

which gives payoffs

$$\pi_1\left(\frac{1}{3}, \frac{1}{3}\right) = \pi_2\left(\frac{1}{3}, \frac{1}{3}\right) = \left(1 - \frac{2}{3}\right)\frac{1}{3} = \frac{1}{9}.$$

- What if player 1 moves first? (Stackelberg model)
- After observing the quantity choice of player 1, player 2 chooses his quantity.
- Given the observed q_1, firm 2 chooses q_2. Optimal choice is

$$BR_2(q_1) = \frac{1-q_1}{2}.$$

- Player 1 should then choose:

$$\max_{q_1} u_1(q_1, BR_2(q_1)) = \left(1 - q_1 - \frac{1-q_1}{2}\right) q_1 = \frac{(1-q_1)q_1}{2}.$$

- This leads to

$$q_1 = \frac{1}{2},\ q_2 = BR_2(q_1) = \frac{1}{4}$$

with payoffs

$$\pi_1\left(\frac{1}{2},\frac{1}{4}\right) = \frac{1}{8},\ \pi_2\left(\frac{1}{2},\frac{1}{4}\right) = \frac{1}{16}.$$

Example: matching pennies

	Head	*Tail*
Head	$1,-1$	$-1,1$
Tail	$-1,1$	$1,-1$

- Nash equilibrium, where both players mix with $1/2$ probabilities
- What if player 1 moves first?
- What if player 1 moves first, but choice is unobservable?

Discussion

- These examples illustrate that the order of moves is crucial
- Moving first may help (commitment to an action)
- Moving first may also hurt (matching pennies)
- The normal form representation misses the dynamic nature of events, so we need to utilize extensive form representation

Sequential rationality

- The key principle is *sequential rationality*, which means that a player should always use a continuation strategy that is optimal given the current situation
- For example, once entrant has entered, the incumbent should act optimally given this fact (accommodate)
- This will lead us to refinements of Nash equilibrium, in particular *subgame perfect equilibrium (SPE)* and *sequential equilibrium*

Subgame

- Consider an extensive form game with perfect recall
- A subgame is a subset of the original game-tree that inherits information sets and payoffs from the original game, and which meets the following requirements:
    1. There is a single node such that all the other nodes are successors of this node (that is, there is a single initial node to the subgame)
    2. Whenever a node belongs to the subgame, then all of its successor nodes must also belong to the subgame.
    3. Whenever a node belongs to the subgame, then all nodes in the same information set must also belong to the subgame.

Subgame perfect equilibrium

Definition

A strategy profile σ of an extensive form game is a subgame perfect Nash equilibrium (SPE) if it induces a Nash equilibrium in every subgame of the original game.

- Every subgame perfect equilibrium is a Nash equilibrium, but the converse is not true. Thus, subgame perfection is a *refinement* of the Nash equilibrium concept.
- The idea is to get rid of equilibria that violate sequential rationality principle.
- It is instructive to go through the examples that we've done so far and identify subgame perfect equilibria.

Backward induction

- In finite games with perfect information, sub-game perfect equilibria are found by backward induction:
- Consider the nodes, whose immediate successors are terminal nodes
- Specify that the player who can move in those nodes chooses the action that leads to the best terminal payoff for her (in case of tie, make an arbitrary selection)
- Then move one step back to the preceding nodes, and specify that the players who move in those nodes choose the action that leads to the best terminal payoff - taking into account the actions specified for the next nodes
- Continue this process until all actions in the game tree have been determined

Theorem

A finite game of perfect information has a subgame perfect Nash equilibrium in pure strategies.

- The proof is the backward induction argument outlined above
- If optimal actions are unique in every node, then there is a unique sub-game perfect equilibrium
- Note that the terminal nodes are needed to start backward induction. Does not work for infinite games.
- How about chess?

Example: chain store paradox

- Note that the entry game (predation) discussed at the beginning of the lecture is a perfect information game
- Could a firm build a reputation for fighting if it faces a sequence of entrants?
- *Chain store paradox* considers an incumbent firm CS that has branches in cities $1, \dots, K$.
- In each city there is a potential entrant.
- In period k, entrant of city k enters or not. If it enters, incumbent may fight or accomodate.

- Payoffs for city k are as in original entry game:

	Fight if entry	Accommodate if entry
Enter	$-1, 0$	$1, 2$
Stay out	$0, 4$	$0, 4$

- Incumbent maximizes the sum of payoffs over all cities, while each entrant maximizes profits of that period.
- An entrant only enters if it knows the CS does not fight.
- Would it pay for CS to build a reputation of toughness if K is very large?

- The paradox is that in SPE, the CS can not build a reputation.
- In the final stage, the optimal action of CS is Accomodate, if the entrant enters.
- The entrant know this, and thus enters.
- By backward induction, this happens in all stages.
- We find that the unique SPE is that all entrants enter and CS always accomodates.
- To bring reputation effects to life, we would need to introduce *incomplete* information (later in this course)

Example: centipede game

- Centipede game is a striking example of backward induction
- Two players take turns to choose Continute (C) or stop (S)
- The game can continue at most K steps (K can be arbitrarily large)
- In stage 1, player 1 decides between C and S. If he chooses S, he gets 2 and player 2 gets 0. Otherwise game goes to stage 2.
- In stage 2, player 2 decides between C and S. If he chooses S, he gets 3 and player 2 gets 1. Otherwise game goes to stage 3, and so on.
- If i stops in stage k, he gets $k+1$, while j gets $k-1$.
- If no player ever stops, both players get K.
- Draw extensive form and solve by backward induction. What is the unique SPE?

Multi-stage games with observed actions

- One restrictive feature of games of perfect information is that only one player moves at a time
- A somewhat larger class of dynamic games is that of multi-stage games with observed actions
- Many players may act simultaneously within each stage
- We may summarize each node that begins stage t by history h^t that contains all actions taken in previous stages: $h^t := \left(a^0, a^1, ..., a^{t-1}\right)$
- A pure strategy is a sequence of maps s_i^t from histories to actions $a_i^t \in A_i\left(h^t\right)$
- Payoff u_i is a function of the terminal history h^{T+1}

One-step deviation principle(1)

- Since many players may act simultaneously within a stage, backward induction argument can not be applied as easily as with games of perfect information
- However, the following principle that extends backward induction idea is useful:

Theorem

In a finite multi-stage game with observed actions, strategy profile s is a subgame perfect equilibrium if and only if there is no player i and no strategy s_i' that agrees with s_i except at a single t and h^t, and such that s_i' is a better response to s_{-i} than s_i conditional on history h^t being reached.

One-step deviation principle(2)

- To check if s is a SPE, we only need to check if any player can improve payoffs by a one-step deviation
- Note that the result requires a finite horizon, just like backward induction
- Some applications have an infinite horizon in which case payoffs defined as functions of the infinite sequence of actions
- Importantly, the result carries over to such games under an extra condition that essentially requires that distant events are relatively unimportant
- In particular, if payoffs are discounted sums of per period payoffs, and payoffs per period are uniformly bounded, then this condition holds
- The proof of the one-step deviation principle is essentially the principle of optimality for dynamic programming.

Example: repeated prisoner's dilemma

- Consider prisoner's dilemma with payoffs:

	Cooperate	*Defect*
Cooperate	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

- Suppose that two players play the game repeatedly for T periods, with total payoff

$$\frac{1-\delta}{1-\delta^T}\sum_{t=0}^{T-1}\delta^t g_i\left(a^t\right),$$

where g_i gives the per-period payoff of action profile a^t as given in the table above

- Players hence maximize their discounted sum of payoffs, where the term $\frac{1-\delta}{1-\delta^T}$ is just a normalization factor to make payoffs of games with different horizons easily comparable
- Suppose first that the game is played just once ($T = 1$). Then (Defect,Defect) is the unique Nash equilibrium (Defect is a dominant action)
- Suppose next that T is finite. Now, subgame perfection requires both players to defect in the last period, and backward induction implies that both players always defect.

- Finally, suppose that T is infinite. Then backward induction cannot be applied but one-step deviation principle holds (discounting and bounded payoffs per period)
- "Both defect every period" is still a SPE
- However, provided that δ is high enough, there are now other SPEs too
- By utilizing one-step deviation principle, show that the following is a SPE: "cooperate in the first period and continue cooperating as long as no player has ever defected. Once one of the players defect, defect in every period for the rest of the game".

Application: Alternating offers bargaining

- We now consider a well known application of dynamic games of perfect information: alternating offers bargaining
- For more on bargaining games, see e.g. Mailath Ch. 9, Osborne and Rubinstein Ch. 7, Fudenberg and Tirole Ch. 4, Myerson Ch. 8
- The original source for the model that we analyze here is Rubinstein (1982), "Perfect equilibrium in a bargaining model", Econometrica 50.

One offer round

- Start with the simplest case: only one offer round
- Then the game collapses to a well-known game called *ultimatum game*
- Two players share a pie of size 1.
- Player 1 suggests a division $x \in (0, 1)$.
- Player 2 accepts or rejects.
- In the former case, 1 gets x and 2 gets $1 - x$. In the latter case, both get 0.

- Given any $x \in (0,1)$, the strategy profile $\{a_1 = x, a_2 = (\text{accept iff } a_1 \leq x)\}$ is a Nash equilibrium. So, there are infinitely many Nash equilibria.
- But once player 1 has made an offer, the optimal strategy for 2 is to accept *any* offer $a_1 < 1$ and he is indifferent with accepting offer $a_1 = 1$.
- What can you say about subgame perfect equilibria?

- Suppose next that after 2 rejects an offer, the roles are changed
- 2 makes an offer x for player 1 and if accepted, she gets $1 - x$ for herself
- What is the SPE of this two-round bargaining game?
- What if players are impatient and payoff in stage 2 is only worth $\delta_i < 1$ times the payoff in stage 1 for player i.
- What is the SPE of this game?

Generalization to longer horizons: Alternating offers bargaining

- The player that rejects an offer makes a counteroffer
- Players discount every round of delay by factor δ_i, $i = 1, 2$
- If there are T periods, we can solve for a SPE by backward induction.
- Suppose we are at the last period, and player 1 makes the offer. Then he should demand the whole pie $x = 1$ and player 2 should accept.

- Suppose we are at period $T-1$, where player 2 makes the offer. He knows that if his offer is not accepted, in the next period player 1 will demand everything. So he should offer the least amount that player 1 would accept, that is $x=\delta_1$.
- Similarly, at period $T-2$ player 1 should offer division $x=1-\delta_2\left(1-\delta_1\right)$, and so on

- Can you show that as $T \to \infty$, then the player i who starts offers
$$x = \frac{1-\delta_j}{1-\delta_i\delta_j}$$
in the first period, and this offer is accepted?
- Note that the more patient player is stronger.

- What if there is infinite horizon? With no end point (i.e. all rejected offers are followed by a new proposal), backward induction is not possible
- Use the concept of SPE
- Notice that the subgame starting after two rejections looks exactly the same as the original game
- Therefore the set of SPE also is the same for the two games

- The famous result proved by Rubinstein (1982) shows that the infinite horizon game also has a unique equilibrium

Theorem

A subgame perfect equilibrium in the infinite horizon alternating offer bargaining game results in immediate acceptance. The unique subgame perfect equilibrium payoff for player 1 is

$$v = \frac{1-\delta_2}{1-\delta_1\delta_2}.$$

Proof of Rubinstein's result (1)

- The part on immediate acceptance is easy and left as an exercise.
- Calculate first the largest subgame perfect equilibrium payoff $\overline{v}$ for player 1 in the game.
- Denote by $v_2(2)$ the continuation payoff to player 2 following a rejection in $T = 1$. The largest payoff for 1 consistent with this continuation payoff to 2 is:

$$1 - \delta_2 v_2(2)$$

- Hence the maximal payoff to 1 results from the minimal $v_2(2)$.

Proof of Rubinstein's result (2)

- We also know that

$$v_2(2) = 1 - \delta_1 v_1(3)$$

- Hence $v_2(2)$ is minimized when $v_1(3)$ is maximized.
- Notice next that the game starting after two rejections is the same game as the original one. Hence $\overline{v}$ is also the maximal value for $v_1(3)$.
- Hence combining the equations, we have

$$\overline{v} = 1 - \delta_2(1 - \delta_1 \overline{v})$$

And hence

$$\overline{v} = \frac{1 - \delta_2}{1 - \delta_1 \delta_2}.$$

Proof of Rubinstein's result (3)

- Denote by $\underline{v}$ the smallest subgame perfect equilibrium payoff to 1. The same argument goes through exchanging everywhere words minimal and maximal. Hence we have:

$$\underline{v} = 1 - \delta_2(1 - \delta_1 \underline{v})$$

and

$$\underline{v} = \frac{1 - \delta_2}{1 - \delta_1 \delta_2}.$$

And thus the result is proved.

Discussion

- If $\delta_1 = \delta_2 = \delta \to 1$, then the SPE payoff converges to 50-50 split
- This is a theory that explains bargaining power by patience
- Cannot explain why there is often delays in bargaining
- Hard to generalize to more than two players
- Must have perfectly divisible offers
- Sensitive to bargaining protocol
- This model is based on Rubinstein (1982), "Perfect equilibrium in a bargaining model", Econometrica 50.

Repeated games

- We now consider repeated games
- This is an important and well studies class of dynamic games
- We restrict to the class of models with "perfect monitoring"
- This falls into the category of multi-stage games with observed actions
  - sub-game perfect equilibrium as solution concept
  - apply one-step deviation principle

Repeated games: literature sources

- In these slides, we only give some basic results and intuitions, and restrict to the case of perfect monitoring (i.e. both players observe perfectly each others' previous actions)
- For more on repeated games, see Mailath Ch. 7, Fudenberg and Tirole Ch. 5, Osborne and Rubinstein Ch. 8, Myerson Ch. 7, Maschler, Solan, and Zamir Ch. 13
- A particularly good and exhaustive source is the book Mailath and Samuelson: Repeated Games and Reputations, 2006, Oxford University Press.

The idea

- In repeated games, the same "stage game" is repeated over and over again
- Player's payoff is most typically the discounted sum of the payoffs across stages
- The underlying idea: players may punish other players' deviations from nice behavior by their future play
- This may discipline behavior in the current period
- As a result, more cooperative behavior is possible

Stage game

- A stage game is a finite I-player simultaneous-move game
- Denote by A_i, $i = 1, ..., I$ the action spaces within a stage
- Stage-game payoff given by

$$g_i : A \to \mathbb{R}.$$

- In an infinite horizon repeated game, the same stage game is repeated forever

Strategies and payoffs

- Players observe each other's actions in previous periods
- Therefore, this is a multi-stage game with observed actions
- Denote by $a^t := (a_1^t, ..., a_I^t)$ the action profile in stage t
- As before, history at stage t, $h^t := \left(a^0, ..., a^{t-1}\right) \in H^t$, summarizes the actions taken in previous stages
- A pure strategy is a sequence of maps s_i^t from histories to actions

- A mixed (behavior) strategy σ_i is a sequence of maps from histories to probability distributions over actions:

$$\sigma_i^t : H^t \to \Delta(A_i).$$

- The payoffs are (normalized) discounted sum of stage payoffs:

$$u_i(\sigma) = \mathbb{E}_\sigma (1-\delta) \sum_{t=0}^{\infty} \delta^t g_i\left(\sigma^t\left(h^t\right)\right),$$

where expectation is taken over possible infinite histories generated by σ

- The term $(1-\delta)$ just normalizes payoffs to "per-period" units
- Note that every period begins a proper subgame
- For any σ and h^t, we can compute the "continuation payoff" at the current stage:

$$\mathbb{E}_\sigma (1-\delta) \sum_{\tau=t}^{\infty} \delta^\tau g_i \left(\sigma^\tau \left(h^\tau\right)\right).$$

- A preliminary result:

Theorem

If $\alpha^* = (\alpha_1^*, ..., \alpha_I^*) \in \Delta(S_1) \times ... \times \Delta(S_I)$ is a Nash equilibrium of the stage game, then the strategy profile

$$\sigma_i^t\left(h^t\right) = \alpha_i^* \text{ for all } i \in I,\ h^t \in H^t,\ t = 0, 1, ...$$

is a sub-game perfect equilibrium of the repeated game. Moreover, if the stage game has m Nash equilibria $\left(\alpha^1, ..., \alpha^m\right)$, then for any map $j(t)$ from time periods to $\{1, ..., m\}$, there is a subgame perfect equilibrium

$$\sigma^t\left(h^t\right) = \alpha^{j(t)},$$

i.e. every player plays according to the stage-game equilibrium $\alpha^{j(t)}$ in stage t.

- Check that you understand why these strategies are sub-game perfect equilibria of the repeated game
- These equilibria are not very interesting. The point in analyzing repeated games is, of course, that more interesting equilibria exist too

Folk theorems

- What kind of payoffs can be supported in equilibrium?
- The main insight of the so-called folk theorems (various versions apply under different conditions) is that virtually any "feasible" and "individually rational" payoff profile can be enforced in an equilibrium, provided that discounting is sufficiently mild

Individually rational payoffs (1)

- What is the lowest payoff that player i's opponents can impose on i?
- Let

$$\underline{v}_i := \min_{\alpha_{-i}} \max_{\alpha_i} g_i\left(\alpha_i, \alpha_{-i}\right),$$

where $\alpha_i \in \Delta\left(S_i\right)$ and $\alpha_{-i} \in \times_{j \neq i} \Delta\left(S_j\right)$

- It is easy to prove the following:

Individually rational payoffs (2)

Theorem

Player i's payoff is at least $\underline{v}_i$ in any Nash equilibrium of the repeated game, regardless of the level of the discount factor.

- Hence, we call $\{(v_1, ..., v_I) : v_i \geq \underline{v}_i \text{ for all } i\}$ the set of individually rational payoffs.

Feasible payoffs (1)

- We want to identify the set of all payoff vectors that result from some feasible strategy profile
- With independent strategies, feasible payoff set is not necessarily convex
  - e.g. in the coordination game with conflicting interest (originally known as "battle of sexes"), payoff $\left(\frac{3}{2}, \frac{3}{2}\right)$ can only be attained by correlated strategies
- Also, with standard mixed strategies, deviations are not perfectly detected (only actions observed, not the actual strategies)
- But in repeated games, convex combinations can be attained by time-varying strategies (if discount factor is large)

Feasible payoffs (2)

- To sidestep this technical issue, we assume here that players can condition their actions on the outcome of a public randomization device in each period
- This allows correlated strategies, where deviations are publicly detected
- Then, the set of feasible payoffs is given by

$$V = co\left\{v : \exists a \in A \text{ such that } g\left(a\right) = v\right\},$$

where co denotes convex hull operator

Folk theorems

- Having defined individually rational and feasible payoffs, we may state the simplest version of a Folk theorem:

Theorem

For every feasible and strictly individually rational payoff vector v (i.e. an element of $\{v \in V : v_i > \underline{v}_i$ for all $i\}$), there exists a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$ there is a Nash equilibrium of the repeated game with payoffs v.

- The proof idea is simple and important: construct strategies where all the players play the stage-game strategies that give payoffs v as long as no player has deviated from this strategy. As soon as one player deviates, other players turn to punishment strategies that "minmax" the deviating player forever after.
- If the players are sufficiently pationt, any finite one-period gain from deviating is outweighed by the loss caused by the punishment, therefore strategies are best-responses (check the details).

- The problem with this theorem is that the Nash equilibrium constructed here is not necessarily sub-game perfect
- The reason is that punishment can be very costly, so once a deviation has occurred, it may not be optimal to carry out the punishment
- However, if the minmax payoff profile itself is a stage-game Nash equilibrium, then the equilibrium is sub-game perfect
- This is the case in repeated Prisoner's dilemma
- The question arises: using less costly punishements, can we generalize the conclusion of the theorem to sub-game perfect equilibria?
- Naturally, we can use some low-payoff stage-game Nash equilibrium profile as a punishment:

Theorem

Let α^* be a stage-game Nash equilibrium with payoff profile e. Then, for any feasible payoff vector with $v_i > e_i$ for every i, there is a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$ there is a sub-game perfect Nash equilibrium of the repeated game with payoffs v.

- The proof is easy and uses the same idea as in the previous theorem, except here one uses Nash equilibrium strategy profile α^* as the punishment to a deviation
- Because the play continues according to a Nash equilibrium even after deviation, this is a sub-game perfect equilibrium
- Note that the conclusion of this Theorem is weaker than in the previous Theorem in the sense that it only covers payoff profiles where each player gets more than in some stage-game Nash equilibrium

- Is it possible to extend the result to cover all individually rational and feasible payoff profiles?
- Fudenberg and Maskin (1986) show that the answer is positive:

Theorem

For any $v \in V$ such that $v_i > \underline{v}_i$ for all i, there is a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$ there is a sub-game perfect Nash equilibrium of the repeated game with payoffs v.

- To be exact, the theorem requires an additional dimensionality condition on the payoff set, see Fudenberg-Tirole book or the original article for full details

Structure of equilibria

- The various folk theorems show that repeated interaction makes cooperation feasible as $\delta \to 1$
- At the same time, they show that the standard equilibrium concepts do little to predict actual play in repated games: the proofs use just one strategy profile that works if δ is large enough
- The set of possible equilibria is large. Is there a systematic way to characterize behavior in equilibrium for a given fixed δ?
- What is the most effective way to punish deviations?

- At the outset, the problem is complicated because the set of potential strategy profiles is very large (what to do after all possible deviations...)
- Abreu (1988) shows that all subgame perfect equilibrium paths can be generated by simple strategy profiles
- "Simple" means that these profiles consists of $I+1$ equilibrium *paths*: the actual play path and I punishment paths.
- A path is just a sequence of action profiles
- This is a relatively simple object - does not contain description of players' behavior after deviations

- The idea is that a deviation is punished by switching to the worst subgame perfect equilibrium path for the deviator:
  - Take a path as a candidate for a subgame perfect equilibrium path. We want to define a simple strategy profile that is a SPE and supports this path.
  - Find the worst sub-game perfect equilibrium path for each player. These are used as "punishment" paths.
  - Define players' behavior: follow the default path as long as no player deviates.
  - If one player deviates, switch to the punishment path of the deviator.

- If there is another deviation from the punishment srategy, again switch to the equilibrium that punishes deviator.
- By one step deviation principle, this is a sub-game perfect equilibrium that replicates the original equilibrium path (recall that one-step deviation principle works for infinite horizon games with discounting)
- Note that once all players follow these strategies, there is no deviation and hence punishment are not used along equilibrium path

- For a formalization of this, see Abreu (1988): "On the theory of infinitely repeated games with discounting", Econometrica 56 (2).
- To illustrate the structure of equilibria, it is very useful to represent strategies as *automata* (see Mailath, ch. 7.1.4)

Example: oligopoly

- Finding the worst possible SPE for each player, as the construction above requires, may be difficult
- However, for symmetric games, finding the worst *strongly symmetric* pure-strategy equilibrium is much easier
- A strategy profile is strongly symmetric, if for all histories h^t and all players i and j, we have $s_i(h^t) = s_j(h^t)$
- Following the same idea as in Abreu (1988), we can construct the best strongly symmetric equilibrium by finding the worst punishment paths in the class of strongly symmetric equilibria
- This works nicely in games where arbitrarily low stage-game payoffs may be induced by symmetric strategies

- As an example, consider a quantity setting oligopoly model (with continuum action spaces)
- This originates from Abreu (1986), "Extremal equilibria of oligopolistic supergames", Journal of Economic Theory (here adapted from Mailath-Samuelson book)
- There are n firms, producing homogeneous output with marginal cost $c < 1$
- Firms maximize discounted sum of stage payoffs with discount factor δ

- Given outputs $q_1, \ldots, q_n$, stage payoff of firm i is

$$u_i \left(q_1, \ldots, q_n\right) = q_i \left(\max \left\{1 -_{j=1}^{n} q_j, 0\right\} - c\right).$$

- The stage game has a unique symmetric Nash equilibrium

$$q_i^N = \frac{1-c}{n+1} := q^N, \; i = 1, \ldots, n$$

with stage payoffs

$$u_i \left(q_1^N, \ldots, q_n^N\right) = \left(\frac{1-c}{n+1}\right)^2.$$

- The symmetric output that maximizes joint profits is

$$q_i^m = \frac{1-c}{2n} := q^m$$

giving payoffs

$$u_i\left(q_1^m, \ldots, q_n^m\right) = \frac{1}{n}\left(\frac{1-c}{2}\right)^2.$$

- Note that in this model, one sub-game perfect equilibrium is trivially $s_i\left(h^t\right) = q^N$ for all i and h^t
- Therefore, if δ is high enough, optimal outputs are achieved by Nash-reversion strategies: play q^m as long as all the players do so, otherwise revert to playing q^N forever

- However, cooperation at lower discount rates is possible with more effective punishments as follows
- Let $\mu(q)$ denote a payoff with symmetric output profile $q_i = q$ for $i = 1, ..., n$:

$$\mu(q) = q\left(\max\{1 - nq, 0\} - c\right)$$

- Let $\mu^d(q)$ denote maximal "deviation payoff" for i when others produce q:

$$\begin{aligned} \mu^d(q) &= \max_{q_1} u_1(q_1, q, ..., q) \\ &= \begin{cases} \frac{1}{4}\left(1 - (n-1)q - c\right)^2 & \text{if } 1 - (n-1)q - c \geq 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

- Note that $\mu(q)$ can be made arbitrarily low with high enough q, allowing severe punishments
- Also, $\mu^d(q)$ is decreasing in q and $\mu^d(q) = 0$ for q high enough
- Let v^* denote the worst payoff achievable in strongly symmetric equilibrium (can be shown as part of the construction that a strategy profile achieving this minimum payoff exists)

- Given this, the best payoff that can be achieved in SPE is obtained by every player choosing q^* given by

$$q^* = \arg\max_q \mu(q) \tag{1}$$

subject to

$$\mu(q) \geq (1-\delta)\,\mu^d(q) + \delta v^*, \tag{2}$$

where the inequality constraint ensures that playinng q^* (now and forever) is better than choosing the best deviation and obaining the worst SPE payoff from that point on

- How do we find v^*?

- The basic insight is that we can obtain v^* by using a "carrot-and-stick" punishment strategy with some "stick" output q^s and "carrot" output q^c
- According to such strategy, choose output q^s in the first period and thereafter play q^c in every period, unless any player deviates from this plan, which causes this prescription to be repeated from the beginning
- Intuitively: q^s leads to painfully low profits (stick), but it has to be suffered once in order for the play to resume to q^c

- To be a SPE, such a strategy must satisfy:
  1. Players don't have an incentive to deviate from "carrot":

$$\begin{aligned} \mu(q^c) &\geq (1-\delta)\,\mu^d(q^c) + \delta\left[(1-\delta)\,\mu(q^s) + \delta\mu(q^c)\right] \text{ or} \\ \mu^d(q^c) - \mu(q^c) &\leq \delta\left(\mu(q^c) - \mu(q^s)\right) \qquad (3) \end{aligned}$$

  2. Players dont' have an incentive to deviate from "stick":

$$\mu^d(q^s) \leq (1-\delta)\,\mu(q^s) + \delta\mu(q^c) \qquad (4)$$

- To find the *optimal* "carrot-and-stick" punishment, we can proceed as follows:
- First, guess that joint optimum q^m can be supported in SPE. If that is the case, then let $q^c = q^m$, and let q^s be the worst "stick" that the players still want to carry out (knowning that this restores play to q^m), ie solve q^s from

$$\mu^d(q^s) = (1-\delta)\mu(q^s) + \delta\mu(q^m).$$

- If

$$\mu^d(q^m) - \mu(q^m) \leq \delta(\mu(q^m) - \mu(q^s)),$$

then no player indeed wants to deviate from q^m, and this carrot-and-stick strategy works giving:

$$v^* = (1-\delta)\mu(q^s) + \delta\mu(q^m)$$

- However, if

$$\mu^d (q^m) - \mu (q^m) > \delta (\mu (q^m) - \mu (q^s)),$$

  then the worst possible punishment is not severe enough, and q^m cannot be implemented
- Then we want to find the lowest $q^c > q^m$ for which there is some q^s such that (3) and (4) hold
- This task is accomplished by finding q^c and q^s that solve those two inequalities as "=" (both "incentive constraints" bind)
- Note that this algorithm gives us the solution to (1) - (2): $q^* = q^c$ and $v^* = (1 - \delta) \mu (q^s) + \delta \mu (q^*)$

- Is something lost by restricting to strongly symmetric punishment strategies? If $v^* = 0$, then clearly there cannot be any better asymmetric punishments (every player guarantees zero by producing zero in every period). Then restricting to strongly symmetric strategies is without loss
- However, if $v^* > 0$, then one could improve by adopting asymmetric punishment strategies
- It can be shown that q^* and v^* are decreasing in discount factor δ, and corresponding stick output q^s is increasing in δ
- That is, higher discount factor improves the achievable stage-payoff by making feasible punishments more severe
- For a high enough discount factor, we have $v^* = 0$ and $q^* = q^m$

Sequential rationality in games of imperfect information

- Recall that a subgame starts with an information set that consists of a *single* node
- But in games of imperfect information, there may be few such nodes
- For example, in Bayesian games, where nature chooses a type for each player, the only subgame is the whole game tree (these games will be analyzed in slide set 3)
- In such situations the refinement of subgame perfect equilibrium has no bite
- To evaluate sequential rationality in information sets with many nodes, we must consider the *beliefs* of the player that chooses her action

Belief system

Definition

A belief system μ assigns for each information set h a probability distribution on the nodes of that information set. In other words, $\mu^h(x) \in [0,1]$ gives a probability of node x in information set h, where $\Sigma_{x \in h} \mu^h(x) = 1$.

- In words, μ^h expresses the beliefs of player $\iota(h)$ on the nodes in h conditional on reaching h.

Sequential rationality

- Let b denote some behavior strategy
- Let $u_i(b|\mu^h)$ be the expected utility of player i given that information set h is reached, given that player i's beliefs with respect to the nodes $x \in h$ is given by μ^h, and given that the strategy profile b is played on all information sets that follow h
- Sequential rationality can now be formally stated:

Sequential rationality

Definition

behavior strategy profile b is sequentially rational (given belief system μ) if for all i and all h such that i moves at h,

$$u_i(b|\mu^h) \geq u_i((b_i', b_{-i})|\mu^h) \text{ for each behavior strategy } b_i' \text{ of } i.$$

- In other words, sequential rationality means expected utility maximization at each h given the beliefs at h and given that all future decisions are taken according to b.

- So far we have said nothing about how beliefs are formed
- Let $\mathbf{a}^x$ be the path of actions that leads from x^0 to x. Define

$$P^b(x) = \prod\{b(a) | a \in \mathbf{a}^x\},$$

and

$$P^b(h) = \sum\{P^b(x) | x \in h\}.$$

- To connect beliefs to strategies, we require that they are obtained from the strategies using Bayes' rule:

$$\mu^h(x) = \frac{P^b(x)}{P^b(h)}, \text{ whenever } P^b(h) > 0.$$

- We have:

Perfect Bayesian Equilibrium

Definition

A Perfect Bayesian Equilibrium (PBE) is a pair (b, μ) such that b is sequentially rational given μ and μ is derived from b using Bayes' rule whenever applicable.

Discussion

- What to do with off-equilibrium pats, i.e. information sets such that $P^b(h) = 0$?
- The version of Perfect Bayesian Equilibrium defined above gives full freedom for choosing those beliefs (this version is called *weak* PBE in MWG)
- Why do off-equilibrium beliefs matter? Because they may induce off-equilibrium actions that in turn influence behavior *on* the equilibrium path
- To make the concept of PBE more useful in applications, additional restrictions for off-equilibrium beliefs have been introduced (see e.g. Fudenberg-Tirole section 8.2, or MWG section 13.C), but this is not a general cure as it may lead to non-existence problems

Sequential equilibrium

The solution concept, introduced in Kreps and Wilson (1982, Econometrica), called *sequential equilibrium* derives beliefs at off-equilibrium information sets as limits from strategies that put a positive but small probability on all actions (so that all information sets are reached with positive probability):

Definition

A pair (b, μ) is a Sequential Equilibrium if:
1) Sequential Rationality: b is sequentially rational given μ
2) Consistency of beliefs: there exists a sequence of pairs $(b^n, \mu^n) \to (b, \mu)$, such that for all n, b^n puts a positive probability on all availabe actions, and for any h and any $x \in h$, $\mu_h^n(x) = P^{b^n}(x)/P^{b^n}(h)$.

Discussion

- Every finite extensive form game with perfect recall has a sequential equilibrium
- In practice, PBE is a popular solution concept in applications
- Sequential equilibrium is important because:
  - Existence is guaranteed (in finite games with perfect recall)
  - Every sequential equilibrium is at the same time a (weak) perfect Bayesian equilibrium
  - Also, if (b, μ) is a sequential equlibrium, then at the same time b is a sub-game perfect equilibrium (this does not necessarily hold for a weak PBE).

Discussion

- A related concept is called extensive form trembling-hand perfect Nash equilibrium, which also always exists in finite games (see MWG Appendix B to Ch. 9). An extensive form trembling-hand perfect equilibrium is a sequential equilibrium, but the converse is not necessarily true.