14.126 GAME THEORY

MIHAI MANEA

Department of Economics, MIT,

1. Forward Induction in Signaling Games

Consider now a **signaling game**. There are two players, a sender S and a receiver R. There is a set T of types for the sender; the realized type will be denoted by t. $p(t)$ denotes the probability of type t. The sender privately observes his type t, then sends a message $m \in M(t)$. The receiver observes the message and chooses an action $a \in A(m)$. Finally both players receive payoffs $u_S(t, m, a), u_R(t, m, a)$; thus the payoffs potentially depend on the true type, the message sent, and the action taken by the receiver.

In such a game we will use $T(m)$ to denote the set $\{t \mid m \in M(t)\}$.

The beer-quiche game from before is an example of such a game. T is the set $\{weak, surly\}$; the messages are $\{beer, quiche\}$; the actions are $\{fight, not\ fight\}$. As we saw before, there are two sequential equilibria: one in which both types of sender choose beer, and another in which both types choose quiche. In each case, the equilibrium is supported by some beliefs such that the sender is likely to have been weak if he chose the unused message, and the receiver responds by fighting in this case.

Cho and Kreps (1987) argued that the equilibrium in which both types choose quiche is unreasonable for the following reason. It does not make any type for the weak type to deviate to ordering beer, no matter how he thinks that the receiver will react, because he is already getting payoff 3 from quiche, whereas he cannot get more than 2 from switching to beer. On the other hand, the surly type can benefit if he thinks that the receiver will react by not fighting. Thus, conditional on seeing beer ordered, the receiver should conclude that the sender is surly and so will not want to fight.

Date: March 2, 2010.

On the other hand, this argument does not rule out the equilibrium in which both types drink beer. In this case, in equilibrium the surly type is getting 3, whereas he gets at most 2 from deviating no matter how the receiver reacts; hence he cannot want to deviate. The weak type, on the other hand, is getting 2, and he can get 3 by switching to quiche if he thinks this will induce the receiver not to fight him. Thus only the weak type would deviate, so the sender's belief (that the receiver is weak if he orders quiche) is reasonable.

Now consider modifying the game by adding an extra option for the receiver: paying a million dollars to the sender. Now the preceding argument doesn't rule out the quiche equilibrium — either type of sender might deviate to beer if he thinks this will induce the receiver to pay him a million dollars. Hence, in order for the argument to go through, we need the additional assumption that the sender cannot expect the receiver to play a bad strategy.

Cho and Kreps formalized this line of reasoning in the **intuitive criterion**, as follows. For any set of types $T' \subseteq T$, write

$$BR(T', m) = \cup_{\mu \mid \mu(T')=1} BR(\mu, m)$$

— the set of strategies that R could reasonably play if he observes m and is sure that the sender's type is in T'. Now with this notation established, consider any sequential equilibrium, and let $u_S^*(t)$ be the equilibrium payoff to a sender of type t. Define

$$\tilde{T}(m) = \{t \mid u_S^*(t) > \max_{a \in BR(T(m),m)} u_S(t, m, a)\}.$$

This is the set of types that do better in equilibrium than they could possibly do by sending m, no matter how R reacts, as long as R is playing a best reply to some belief. We then say that the proposed equilibrium fails the intuitive criterion if there exist a type t' and a message m such that

$$u_S^*(t') < \min_{a \in BR(T(m) \setminus \tilde{T}(m), m)} u_S(t', m, a).$$

In words, the equilibrium fails the intuitive criterion if some type t' of the sender is getting a lower payoff than any payoff he could possibly get by playing m if he could thereby convince the sender that he could not possibly be in $\tilde{T}(m)$.

In the beer-quiche example, the all-quiche equilibirum fails this criterion: let $t' = surly$ and $m = beer$; check that $\tilde{T}(m) = \{weak\}$.

Now we can apply this procedure repeatedly, giving the **iterated intuitive criterion**. We can use the intuitive criterion as above to rule out some pairs (t, m) — type t cannot conceivably send message m. Now we can rule out some actions of the receiver, by requiring that the receiver should be playing a best reply to some belief about the types that have not yet been eliminated (given the message). Given this elimination, we can go back and possibly rule out more pairs (t, m), and so forth.

This idea has been further developed by Banks and Sobel (1987). They say that type t' is **infinitely more likely** to choose the out-of-equilibrium message m than type t if the set of best-replies by the receiver that make t' willing to deviate to m is a strict superset of the best-replies that make t willing to deviate. Conditional on observing m, the receiver should put belief 0 on type t in this case. As above, we can apply this to eliminate possible actions by the receiver, and proceed iteratively. This leads to the equilibrium refinement criterion of **universal divinity**.

Banks and Sobel also give a further criterion D2, which says [oops, sorry I missed out on this one]. The motivating application is Spence's job-market signaling model. With just two types of job applicant, the intuitive criterion selects the equilibrium where the low type gets the lowest level of education and the high type gets just enough education to deter the low type. With more types, the intuitive criterion no longer accomplishes this. Universal divinity does manage to uniquely select the separating equilibrium that minimizes social waste by having each type get just enough education to deter the next-lower type from imitating him.

2. Forward Induction in General

The preceding ideas are all attempts to capture some kind of **forward induction**: players should believe in the rationality of their opponents, even after observing a deviation; thus if you observe an out-of-equilibrium action being played, you should believe that your opponent expected you to play in a way that made his action reasonable, and this in turn is informative about his type (or, in more general extensive forms, about how he plans to play in the future).

Consider now the extensive-form game as follows: 1 can play O, leading to $(2, 2)$, or I, leading to the battle-of-the-sexes game There is an SPE in which player 1 first plays O; conditional on playing I, they play the equilibrium (W, T). But the following forward-induction argument suggests this equilibrium is unreasonable: if player 1 plays I, this suggests he is

	T	W
T	$0,0$	$3,1$
W	$1,3$	$0,0$

expecting to coordinate on (T, W) in the battle-of-the-sexes game, so player 2, anticipating this, will play W. Thus if 1 can convince 2 to play W by playing I in the first stage, he can get the higher payoff $(3, 1)$.

This can also be represented in (reduced) normal form.

	T	W
O	$2,2$	$2,2$
IT	$0,0$	$3,1$
IW	$1,3$	$0,0$

This representation of the game shows a connection between forward induction and strict dominance. We can rule out IW because it is dominated by O; then the only perfect equilibrium of the remaining game is (IT, W) giving payoffs $(3, 1)$. However, (O, T) can be enforced as a perfect (in fact a proper) equilibrium in the normal-form game.

Based on this example, Kohlberg and Mertens (1986) define the notion of **stable equilibria**. It is a set-valued concept — not a property of individual equilibrium but of sets of strategies, one for each player. They first argue that their solution concept should meet the following requirements:

- Iterated dominance: every strategically stable set must contain a strategically stable set of any game obtained by deleting a strictly dominated strategy.
- Admissibility: no mixed strategy in a strategically stable set assigns positive probability to a strictly dominated strategy.
- Invariance to extensive-form representation: they define an equivalence relation between extensive forms and require that any stable set in one game should be stable in any equivalent game.

They then define strategic stability in a way such that these criteria are satisfied. Their definition is as follows: A closed set S of NE is **strategically stable** if it is minimal among sets with the following property: for every $\eta > 0$, there exists $\varepsilon' > 0$ such that, for all $\varepsilon < \varepsilon'$, all choices of $0 < \varepsilon(s_i) \leq \varepsilon$ for each player i and strategies s_i, the game where each player i is constrained to play every s_i with probability at least $\varepsilon(s_i)$ has a Nash equilibrium which is within distance η of some equilibrium in S.

Thus, any sequence of ε-perturbed games as $\varepsilon \to 0$ should have equilibria corresponding to an equilibrium in S. Notice that we need the minimality property of S to give bite to this definition — otherwise, by upper hemi-continuity, we know that the set of all Nash equilibria would be strategically stable, and we get no refinement.

The difference with trembling-hand perfection is that there should be convergence to one of the selected equilibria for *any* sequence of perturbations, not just some sequence of perturbations.

They have a theorem that there exists some stable set that is contained in a connected component of the set of Nash equilibria. Generically, each component of the set of Nash equilibria leads to a single distribution over outcomes in equilibrium; thus, generically, there exists a stable set that determines a unique outcome distribution. Moreover, any stable set contains a stable set of the game obtained by elimination of a weakly dominated strategy.

Moreover, stable sets have an even stronger property, "never a weak best reply"; cf. p. 445 of FT. This is meant to be a way of capturing the idea of forward induction.

There are actually a lot of stability concepts in the literature. Mertens has more papers with alternative definitions.

Every equilibrium in a stable set has to be a perfect equilibrium. This follows from the minimality condition — if an equilbrium is not a limiting equilibrium along some sequence of trembles, then there's no need to include it in the stable set. But notice, these equilibria are only guaranteed to be perfect in the normal form, not in the agent-normal form (if the game represented is an extensive-form one).

Some recent papers further develop these ideas. Battigalli and Siniscalchi (2002) are interested in the epistemic conditions that lead to forward induction. They have an epistemic model, with state of nature of the form $\omega = (s_i, t_i)_{i \in N}$, where s_i represents player i's disposition to act and t_i represents his disposition to believe. t_i specifies a belief $g_{i,h} \in \Delta(\Omega_{-i})$ over states of the other players for each information set h of player i. We saw i is rational at state ω if s_i is a best reply to his beliefs t_i at each information set. Let R be the set of states at which every player is rational. For any event $E \subseteq \Omega$, we can define the set

$B_{i,h}(E) = \{(s,t) \in \Omega \mid g_{i,h}(E) = 1\}$, i.e. the set of states where i is sure that E has occurred (at information set h). We can define $B_h(E) = \cap_i B_{i,h}$. Finally $SB_i(E) = \cap_h B_{i,h}(E)$, the set of states at which i **strongly believes** in event E, meaning the set of states at which i would be sure of E as long as he's reached an information set where E is possible. Finally, they show that $SB(R)$ identifies forward induction — that is, in the states of the world where everyone strongly believes that everyone is sequentially rational, strategies must form a profile that is not ruled out by forward induction.

Battigalli and Siniscalchi take this a level further by iterating the strong-beliefs operator — everyone strongly believes that everyone strongly believes that everyone is rational, and so forth — and this operator leads to backward induction in games of perfect information; without perfect information, it leads to iterated deletion of strategies that are never a best reply. This gives a formalization of the idea of rationalizability in extensive-form games.

3. Repeated games

We now consider the standard model of repeated games. Let $G = (N, A, u)$ be a normal-form stage game. At time $t = 0, 1, ldots$, the players simultaneously play game G. At each period, the players can all observe play in each previous period; the history is denoted $h^t = (a^0, \dots, a^{t-1})$. Payoffs in the repeated game $RG(\delta)$ are given by $U_i = (1-\delta_i)\sum_{t=0}^{\infty} \delta_i^t u_i(a^t)$. The $(1-\delta_i)$ factor normalizes the sum so that payoffs in the repeated game are on the same scale as in the stage game. We assume players play behavior strategies (by Kuhn's theorem), so a strategy for player i is given by a choice of $\sigma_i(h^t) \in \Delta(A_i)$ for each history h^t.

Given such strategies, we can define continuation payoffs after any history h^t: $U_i(\sigma|h^t)$.

If α^* is a Nash equilibrium of the static game, then playing α^* at every history is a subgame-perfect equilibrium of the repeated game. Conversely: for any finite game G and any $\varepsilon > 0$, there exists $\bar{\delta}$ with the property that, for any $\delta < \bar{\delta}$, any SPE of the repeated game $RG(\delta)$ has the property that, at every history, play is within ε of a static NE. However, we usually care about players with high discount factors, not low discount factors.

The main results for repeated games are "Folk Theorems": for high enough δ, every feasible and individually rational payoff in the stage game can be enforced as an equilibrium of the repeated game. There are several versions of such a theorem, which is why we use the plural. For now, we look at repeated games with perfect monitoring (as just defined), where the appropriate equilibrium concept is SPE. The way to check an SPE is via the one-shot deviation principle. Payoffs from playing a at history h^t are given by the value function

$$V_i(a) = (1-\delta)u_i(a) + \delta U_i(\sigma|h^t, a). \tag{3.1}$$

This gives us an easy way to check whether or not a player wants to deviate from a proposed strategy, given other player's strategies. σ is an SPE if and only if, for every history h^t, $\sigma|h^t$ is a NE of the induced game $G(h^t, \sigma)$ whose payoffs are given by (3.1).

To state a folk theorem, we need to explain the terms "individually rational" and "feasible." The **minmax payoff** of player i is the worst payoff his opponents can hold him down to if he knows their strategies:

$$\underline{v}_i = \min_{\alpha_{-i} \in \Delta(A_{-i})} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}).$$

We will let m^i, a **minmax profile for** i, denote a profile of strategies (a_i, α_{-i}) (which can involve correlation for players $-i$) that solves this minimization and maximization problem.

In any SPE — in fact, any Nash equilibrium — i's payoff is at least his minmax payoff, since he can always get at least this much by just best-responding to his opponents' actions in each period separately. This motivates us to say that a payoff vector v is **individually rational** if $v_i \geq \underline{v}_i$ for each i, and it is **strictly individually rational** if the inequality is strict for each i.

The set of **feasible payoffs** is the convex hull of the set $\{u(a) \mid a \in A\}$. Again note that this can include payoffs that are not obtainable in the stage game using mixed strategies, because correlation between players may be required.

Also, in studying repeated games we usually assume the availability of a **public randomization device** that produces a publicly observed signal $\omega^t \in [0, 1]$, uniformly distributed and independent across periods, so that players can condition their actions on the signal. Properly, we should include the signals (or at least the current period's signal) in the specification of the history, but it is conventional not to write it out explicitly. (Fudenberg and Maskin (1991) showed that one can actually get rid of the public randomization device for sufficiently high δ, by appropriate choice of which periods to play each action profile involved.)

An easy folk theorem is that of Friedman (1971):

Theorem 1. *If e the payoff vector of some Nash equilibrium of G, and v is a feasible payoff vector with $v_i > e_i$ for each i, then for all sufficiently high δ, there exists an SPE with payoffs v.*

Proof. Just specify that the players play whichever action profile gives payoffs v (using the public randomization device to correlate their actions if necessary), and revert to the static Nash permanently if anyone has ever deviated. □

So, for example, if there is a Nash equilibrium that gives everyone their minmax payoff (for example, in the prisoner's dilemma), then every individually rational and feasible payoff vector is obtainable in SPE. But a more general folk theorem would say that every individually rational, feasible payoff is achievable in SPE under more general conditions. This is harder to show, because in order for one player to be punished by minmax if he deviates, others need to be willing to punish him. Thus, for example, if all players have equal payoffs, then it may not be possible to punish a player for deviating, because the punisher hurts himself as well as the deviator.

For this reason, the standard folk theorem (due to Fudenberg and Maskin, 1986) requires a full-dimensionality condition.

Theorem 2. *Suppose the set of feasible payoffs V has full dimension n. For any feasible and strictly individually rational payoff vector v, there exists $\underline{\delta}$ such that whenever $\delta > \underline{\delta}$, there exists an SPE of $RG(\delta)$ with payoffs v.*

Actually we don't quite need the full-dimensionality condition — all we need, conceptually, is that there are no two players who have the same payoff functions; more precisely, no player's payoff function can be a positive affine transformation of any other's (Abreu, Dutta, & Smith, 1994).

Proof. Assume first that i's minmax action profile m^i is pure. Consider the action profile a for which $u(a) = v$. Choose v' in the interior of the feasible, individually rational set with $v'_i < v_i$ for each i. We can do this by full-dimensionality. Let w^i denote v'_i with ε added to each player's payoff except for player i; with ε low enough, this will again be a feasible payoff vector.

Strategies are now specified as follows.

- Phase I: play a, as long as there are no deviations. If i deviates, switch to II_i.
- Phase II_i: play m^i. If player j deviates, switch to II_j. Note that if m^i is a pure strategy profile it is clear what we mean by j deviating. If it requires mixing it is not so clear; this will be dealt with in the second part of the proof. Phase II_i lasts for N periods, where N is a number to be determined, and if there are no deviations during this time, play switches to III_i.
- Phase III_i: play the action profile leading to payoffs w^i forever. If j deviates, go to II_j. (This is the "reward" phase that gives players $-i$ incentives to punish in phase II_i.)

We check that there are no incentives to deviate, using the one-shot deviation principle for each of the three phases: calculate the payoff to i from complying and from deviating in each phase. Phases II_i and II_j ($j \neq i$) need to be considered separately, as do III_i and III_j.

- Phase I: deviating gives at most $(1-\delta)M + \delta(1-\delta^N)\underline{v}_i + \delta^{N+1}v'_i$, where M is some upper bound on all of i's feasible payoffs, and complying gives v_i. Whatever N we have chosen, it is clear that as long as δ is sufficiently close to 1, complying produces a higher payoff than deviating, since $v'_i < v_i$.
- Phase II_i: Suppose there are $N' \leq N$ remaining periods in this phase. Then complying gives i a payoff of $(1-\delta^{N'})\underline{v}_i + \delta^{N'}v'_i$, whereas since i is being minmaxed, deviating can't help in the current period and leads to N more periods of punishment, for a total payoff of at most $(1-\delta^{N+1})\underline{v}_i + \delta^{N+1}v'_i$. Thus deviating is always worse than complying.
- Phase II_j: With N' remaining periods, i gets $(1-\delta^{N'})u_i(m^j) + \delta^{N'}(v_j + \varepsilon)$ from complying and at most $(1-\delta)M + (\delta - \delta^N)\underline{v}_i + \delta^N v'_i$ from deviating. When δ is large enough, complying is preferred.
- Phase III_i: This is the one case that affects the choice of N. Complying gives v'_i in every period, while deviating gives at most $(1-\delta)M + \delta(1-\delta^N)\underline{v}_i + \delta^{N+1}v'_i$. Canceling out common terms, the comparison is between $((1-\delta^{N+1})/(1-\delta))v'_i$ and $M + ((1-\delta^N)/(1-\delta))\underline{v}_i$. The fractions approach $N+1$ and N as $\delta \to 1$. So for sufficiently large N and δ close enough to 1 the desired inequality will hold.

- Phase III_j: [fill in]

Now we need to deal with the part where minmax strategies are mixed. This involves having the continuations use path-dependent ε rewards so that every player is indifferent among all the actions in the support of his minmax strategy — we'll talk about it more next time. $\square$