An Infinite Extensive Form

Debraj Ray, September 2006

1. Model

I describe an infinite extensive form which captures a large variety of economic situations.

The ambient action sets for individual i will be denoted by a sequence $A_i(t)$, each a compact metric space with generic action $a_i(t)$. Let $\mathbf{a}(t)$ be the action *profile*, belonging to the product action set $A(t)$. We assume that all action profiles are commonly observed and remembered.

The game starts at date 0. A *path* is a sequence of action profiles $\mathbf{a} = \{\mathbf{a}(t)\}$, such that each $\mathbf{a}(t) \in A(t)$.

All paths won't be feasible, as individual choices may be restricted by the history of the game so far. This is what we turn to next.

A *history* at date $t \geq 1$ (or a t-history) is a full description of all that has transpired in the past, up to and including date $t-1$. Obviously, a t-history may be identified with a finite sequence of action profiles up to $t-1$. To streamline the definitions, introduce an arbitrary singleton $h(0)$ which will serve as a (the) 0-history. Let $H(t)$ be the set of all (feasible) t-histories, to be defined recursively below.

For all i and $t \geq 0$, there exists a *feasibility correspondence* $C_i(t, .)$ that maps elements in $H(t)$ to subsets of $A_i(t)$, to collections of *feasible actions*. By convention, $C_i(0, h(0)) = A_i(0)$ for all i.

Now we can inductively define the set of all feasible t-histories. At date 0 $H(0)$ is just a singleton set. Recursively, if $H(t)$ has already been defined, then the set of all $t+1$-histories is the collection

$$H(t+1) = \{(h(t), a(t)) | h(t) \in H(t) \text{ and } a_i(t) \in C_i(t, h(t)) \text{ for all } i\}.$$

A *feasible path* is a sequence $\mathbf{a}$ of action profiles with "initial segments" that are feasible t-histories for all t. Specifically, $\mathbf{a}$ is feasible if the *truncations* $(h(0), \mathbf{a}(0), \ldots \mathbf{a}(t)) \in H(t+1)$ for all $t \geq 0$.[1] Let H be the set of all feasible paths.

[C.1] For each i and date t, $C_i(t, .)$ is a nonempty-valued, continuous correspondence.

Recall that H can be viewed as a subset of $\bigtimes_t A(t)$. Endowed with the product topology, this latter set is obviously compact. This allows us to establish

Lemma 1. *Under [C.1], H is compact in the product topology.*

Proof. Simply prove that H is closed. A sequential argument is very easy to provide. Let $\mathbf{a}^k$ be a sequence of paths in H converging in the product topology (so pointwise) to some sequence $\mathbf{a} \in \bigtimes_t A(t)$. We must show that $\mathbf{a} \in H$ as well. Do so inductively. Suppose that for some t, the truncation $h(t)$ of $\mathbf{a}$ lies in $H(t)$. (This is trivially true at $t = 0$.) Consider the corresponding

[1] Note that we need to slip in the arbitrary singleton $h(0)$ to keep the notation consistent.

truncations $h^k(t)$ (of $\mathbf{a}^k$) that converge pointwise to $h(t)$. By definition, $a_i^k(t) \in C_i(t, h_i^k(t)$ for every k and i. Because C_i is continuous, this must be true at the pointwise limit as well. So $\mathbf{a}(t)$ has the property that $a_i(t) \in C_i(t, h_i(t)$ for every i, which means that the truncation $h(t+1)$ lies in $H(t+1)$. This recursion proves that $\mathbf{a}$ is a feasible path. □

Note: to do this you dont't need the full power of [C.1].

Finally, we define payoff functions F_i for each individual i, which are real-valued mappings defined on feasible paths H. We assume

[F.1] For each i, F_i is a continuous function on H.

This is effectively the discounting assumption. Notice that the product topology is weak, so the continuity assumption is correspondingly strong. As an example of a well-defined payoff function which is not continuous, look at the cake-eating problem without discounting.

2. Strategies and Equilibrium

A *strategy* for player i is a specification of a feasible action conditional on every t-history. More formally, a strategy σ_i is a collection $\{\sigma_i(t, .)\}$ such that for every $t \geq 0$ and $h(t) \in H(t)$, $\sigma_i(t, h(t)) \in C_i(t, h(t))$.

A *strategy profile* is a collection σ of strategies, one for each player. Such a profile generates a feasible path $\mathbf{a}(\sigma)$. More generally, σ generates a path $\mathbf{a}(\sigma, h(t))$ conditional on every t-history, for $t \geq 0$. In the latter case, the path obviously coincides with the t-history in question up to date $t-1$.

A strategy profile σ is a *Nash equilibrium* if for every player i and every strategy σ_i',

$$F_i(\mathbf{a}(\sigma)) \geq F_i(\mathbf{a}(\sigma_{-i}, \sigma_i')).$$

And σ is a *subgame perfect Nash equilibrium* (SGPE) if for every history $h(t)$, every player i, and every alternative strategy σ_i',

$$F_i(\mathbf{a}(\sigma, h(t))) \geq F_i(\mathbf{a}(\sigma_{-i}, \sigma_i', h(t))).$$

3. Special Cases

3.1. Finite-Horizon Games.

3.1. Finite-Horizon Games. If the horizon is T, simply set $A_i(t)$ equal to a singleton for all i and all $t \geq T$.

3.2. Games of Perfect Information. At every date t and for every t-history $h(t)$, *at most one* set $C_i(t, h(t))$ is not a singleton. Note: it is possible that different players can move at the same date, depending on the particular history.

3.3. **Repeated Games With Discounting.** $A_i(t) = A_i$ for all t, and $C_i(t, h(t)) = A_i$ for every t-history $h(t)$. Moreover, for every i there is a one-period utility indicator $f_i : \bigtimes_j A_j \mapsto \mathbb{R}$ and a discount factor $\delta_i \in (0, 1)$ such that

$$F_i(\mathbf{a}) = (1 - \delta_i) \sum_{t=0}^{\infty} \delta_i^t f_i(\mathbf{a}_t).$$

3.4. **Games with Termination Threats.** A principal offers a contract at each date to an agent, who can comply or deviate. If he complies, the relationship continues. If he deviates, the relationship is over.

3.5. **Games Against Different Selves.** The Strotz model and its later descendants — such as hyperbolic discounting — can be modelled as special cases of the structure here.

3.6. **Games with State Variables.** Consider an oil cartel selling oil on the world market. The state variable is the existing vector of oil stocks. The set of actions at each date is the extent of oil extraction, which is obviously affected by the size of the stock. Build payoff functions as in Cournot oligopoly.

4. The One-Shot Deviation Principle

The one-shot deviation principle is fundamental to the theory of extensive games. It was originally formulated by David Blackwell (1965) in the context of dynamic programming. As the strategy of other players induces a normal maximization problem for any one player, we can formulate the principle in the context of a single-person decision tree.

Fix a strategy σ_{-i} for all agents other than i. This induces a "single player game" with exactly the same notation as above, in which all nonsingleton feasible sets are controlled by i. In what follows, we consider this single-player game, which may be viewed as an optimization problem.

A strategy σ_i is a *perfect best response* if there is no strategy σ_i' and t-history $h(t)$ such that

$$F(\mathbf{a}(\sigma', h(t))) > F(\mathbf{a}(\sigma, h(t))),$$

where $\sigma = (\sigma_i, \sigma_{-i})$ and $\sigma' = (\sigma_i', \sigma_{-i})$.

For any strategy σ, t-history $h(t)$, and any action $a \in C_i(t, h(t))$, use the notation σ^a to denote the strategy obtained by simply substituting the choice a at $h(t)$, instead of the action prescribed there by σ, and leaving all else unchanged. A strategy σ_i is *unimprovable* for i if there is no t-history $h(t)$ and corresponding σ_i^a such that $F_i(\mathbf{a}(\sigma^a, h(t))) > F_i(\mathbf{a}(\sigma, h(t)))$, where $\sigma^a = (\sigma_i^a, \sigma_{-i})$.

Observe that σ_i^a is a special strategy, differing as it does from σ_i by only "a one-shot deviation" at the t-history x. It is therefore obvious that an optimal strategy is unimprovable. The converse is what we're after:

Theorem **1.** *Under [F.1], an unimprovable strategy must be a best response.*

To see why [F.1] is needed, consider the following (rather silly, but perfectly legitimate) one-person game, with an infinite number of nodes. At each node the person can choose a or b. If she chooses a infinitely many times, she gets a payoff of 1, otherwise she gets a payoff of 0. Consider the strategy in which the person chooses b each time she moves. This strategy is sub-optimal, but there is no one-shot profitable deviation from it.

Proof of Theorem 1. Suppose that σ_i is not a best response. Then there exists σ_i' and t-history $h(t)$ such that $F_i(\mathbf{a}(\sigma', h(t))) > F(\mathbf{a}(\sigma_i, h(t)))$. This is equivalent to the following assertion: there is a feasible path $\mathbf{a}$ which coincides with $h(t)$ up to date $t-1$ such that

$$F_i(\mathbf{a}) \geq F_i(\mathbf{a}(\sigma_i, h(t))) + 2\epsilon$$

for some $\epsilon > 0$ Now using [C.1], choose an integer M such that if any feasible path $\mathbf{a}'$ shares the first $(M+t)$-histories as $\mathbf{a}$ (but is free to be quite different thereafter),

$$F_i(\mathbf{a}') \geq F_i(\mathbf{a}) - \epsilon.$$

For all such paths $\mathbf{a}'$, it follows from the two inequalities above that

$$F_i(\mathbf{a}') \geq F(\sigma, h(t)) + \epsilon.$$

In particular, this means that a *finite* number M of one-shot deviations, starting at the t-history $h(t)$, *with σ applied everywhere else*, is enough to generate a payoff improvement for i.

But then at least *one* of these one-shot deviations, applied *alone*, must yield a payoff improvement at (at least) one of the s-histories $h(s)$, for $s = \{t, \ldots, t+N\}$. This proves that if σ_i is not a best response, it must be improvable. □

5. Existence Questions

The backward induction argument runs into different sorts of problems in infinite models. This happens even if the horizon is finite.

This example is from Hellwig, Leininger, Reny and Robson (1990). Assume there are two players, 1 and 2, moving sequentially. Player 1 chooses a_1 from the set $[-1, 1]$, then player 2 chooses a_2 from the set $[-1, 2]$. Player 1's payoff function is $a_1 - a_2$. Player 2's payoff function is $a_1 a_2$.

Proceeding by backward induction, it is obvious that 2's strategy must be a selection from the correspondence

$$\begin{aligned} C(a_1) &= \{-1\} \text{ if } a_1 < 0 \\ &= [-1, 2] \text{ if } a_1 = 0 \\ &= \{2\} \text{ if } a_1 > 0. \end{aligned}$$

But now notice that most selections from C will make things difficult for player 1. For instance, the strategy selection that selects 2 at the "jump" will leave player 1 with no best response! Indeed, the only selection that does leave player 1 with a best response is the one that selects -1 at the jump. Then player 1 has a well-defined best response.

So this selection has to be done carefully if existence needs to be guaranteed. But just how is the selection to be made? One possibility is that we do it to make sure that the immediate predecessor in the sequential moves game is put in the best possible position, then select again and so forth. But this is problematic: just introduce a dummy player in the above example who moves in between.

On another matter, the selection may have to be contingent on the full history, so that no Markovian equilibrium might exist even when the underlying model is Markovian. The most definite instance of this is the Peleg-Yaari example (1973), done in the context of the Strotz model of changing tastes (1956). The problem is simple: there are four periods over which to eat a cake — $c(t)$, $t = 0, 1, 2, 3$ — and four persons (or personalities) in charge of the four consumptions. Let $y(t)$ be the stock of uneaten cake at the start of date t; then $y(t+1) = y(t) - c(t)$. Let

$$\begin{aligned} f_3(\mathbf{c}) &= c_3 \\ f_2(\mathbf{c}) &= \min\{2c(2), \frac{c(2)+3}{2}\} + c(3) \\ f_1(\mathbf{c}) &= \min\{2c(1), \frac{c(1)+3}{2}\} + c(3) \\ f_0(\mathbf{c}) &= [c(0)c(1)c(2)]^{1/3} + c(1) \end{aligned}$$

Try backward induction. Then the last person (3) eats of all of the cake. Person 2 will eat all the cake up to 1 unit and bequeath any excess, so his consumption strategy is unique and conditioned only on $y(2)$; it is

$$\begin{aligned} c_2(y) &= y \text{ for } y \leq 1 \\ &= 1 \text{ for } y > 1. \end{aligned}$$

Now consider 1. She does not care about 2's consumption, so it is easy to see that she will consume everything until $y(1) = 3$ after which she will consume 1 and bequeath the excess. Her optimal correspondence is as follows:

$$\begin{aligned} c_1(y) &= \{y\} \text{ for } y < 3 \\ &= \{1, 3\} \text{ for } y = 3 \\ &= \{1\} \text{ for } y > 3. \end{aligned}$$

Now if player 0 is allowed to choose from this correspondence (as in the previous example), he will still have a maximum for every initial condition, *but the selection will vary with the player's initial condition*. To see this, note that if the selection ascribes invariant probability p to the choice of 1 when $y(1) = 3$, then

$$\begin{aligned} f_0 &= c(0)^{1/3} + 1 \text{ for } 0 \leq c(0) < y(0) - 3 \\ &= p[c(0)^{1/3} + 1] + 3(1-p) \text{ for } c(0) = y(0) - 3 \\ &= y(0) - c(0) \text{ for } y(0) - 3 < c(0) \leq y(0). \end{aligned}$$

If $p < 1$, this function fails to reach a maximum whenever $y(0) > 11$. But if $p > 0$ this function fails to reach a maximum whenever $3 \leq y(0) < 11$!

So the conditioning has to be more subtle. It can be shown — see Fudenberg and Levine (1983), Harris (1985) and Hellwig and Leininger (1988) for different cases and levels of generality — that perfect information games nonetheless admit equilibria in history-dependent strategies. Bernheim and Ray (1986) show that Markov equilibria can be obtained if one allows for "small" degrees of uncertainty in the transmission of endowments. By suitable truncation arguments, these arguments apply also to the infinite horizon.