Economics 8117-8 **Noncooperative Game Theory**

March 31, 1998 **Lecture 15**

Professor Andrew McLennan

Games of Incomplete Information

I. Introduction.

A. We have seen many games of incomplete information; we now study them formally.

B. Finite games of incomplete information are now regarded as the central object of noncooperative game theory.

1. Games of perfect information are well understood.
2. Finite games of incomplete information are almost completely general, the principal losses of generality arising from the various types of finiteness.
   a. The information sets are finite.
   b. The sets of actions are finite.
   c. There are finitely many stages of play.
3. The paper of Fudenberg and Levine shows that in many cases results for finite games can be extended to infinite games immediately.
4. Dropping the finiteness assumptions would entangle topological and measure–theoretic complications with the still unresolved question of what are good solution concepts.

C. Outline.

1. Formal description of the class of games.
2. Derived normal forms.
3. Kuhn's Theorem.

II. Notation.

A. As was the case for games of perfect information, the physical possibilities are represented by a pair $(T, \prec)$ where T is a finite set of nodes and $\prec$ is a strict partial ordering of T representing precedence.

1. $P(t) = \{x | x \prec t\}$ is the set of *predecessors* of t.
2. We assume that $(T, \prec)$ is an *arborescence*: $P(t)$ is completely ordered by $\prec$ for all t.
3. $W = \{w \in T | P(w) = \emptyset\}$ is the set of *initial nodes.*
   a. For our purposes it is sufficient to adopt the point of view that all moves of nature that might occur are "decided" before the beginning of the game and are reflected in the choice of $w \in W$.
   b. For card games this means that the deck is shuffled once, and each new card is drawn from the top. The alternative would be to reshuffle the remaining cards before each new card is drawn.
4. $Y = T - W$ is the set of *noninitial nodes.*
5. For $y \in Y$, $p_1(y) = \max P(y)$ is the *immediate predecessor* of y.
6. Adopting the convention that $p_1(w) = w$ for $w \in W$, we can define the ℓ^{th} *predecessor* of y by the formula $p_\ell(y) = p_1(p_{\ell-1}(y))$. We adopt the convention that $p_0(t) = t$.
7. The *level* of t is the smallest integer $\ell(t)$ such that $p_{\ell(t)}(t) = w$.
8. $F(t) = p_1^{-1}(t)$ is the set of *immediate successors* of t.
9. $Z = \{z \in T | F(z) = \emptyset\}$ is the set of *terminal nodes.*
10. $X = T - Z$ is the set of *strategic* or *nonterminal nodes.*
11. For $x \in X$, $Z(x) = \{z \in Z | x \prec z\}$ is the set of *terminal successors* of x.

B. Actions.

1. A is a finite set of *actions* or *moves.*
2. The function $\alpha : Y \to A$ assigns the last action $\alpha(y)$ prior to the occurrence of y to each noninitial node y.
3. **Assumption 1:** If $y, y' \in F(x)$ are distinct, then $\alpha(y) \neq \alpha(y')$. In words the meaning of choosing an action at x must be unambiguous.

C. Information sets.

1. The set of *information sets* is H, a partition of X. To be a *partition* the following conditions must be satisfied.

   a. $\emptyset \neq h \subset X$ for all $h \in H$.

   b. If $h, h' \in H$ are distinct then $h \cap h' = \emptyset$.

   c. $\bigcup_{h \in H} h = X$.

2. For $x \in X$, $H(x)$ is the information set containing x.

3. **Assumption 2:** If $x' \in H(x)$ then $\alpha(F(x)) = \alpha(F(x'))$. In words, the set of available actions is the same at all nodes in an information set, which of course is necessary if the agent is to not know the node.

4. The *set of available actions* at h is $A(h) = \alpha(F(x))$.

5. **Assumption 3:** $\{A(h)|h \in H\}$ is a partition of A.

   a. Basically this is a mathematical convenience.

      i. It does not matter if two actions at different information sets have the same name, either intuitively of formally, since the two actions have different consequences.

      ii. There is no need for actions that are not available at any information set, but neither is there any need to assume that there are none.

D. Agents.

1. $I = \{1, \ldots, n\}$ is the set of *agents.*

2. The agent whose turn it is to move is given by the function $\iota : H \to I$.

3. Let $H^i = \iota^{-1}(i)$ be the set of information sets at which i moves, and let $A^i = \bigcup_{h \in H^i} A(h)$ be the set of all actions that i might ever choose.

E. Perfect Recall

1. First we consider an example without perfect recall.

2. **Assumption 4:** (Perfect Recall) If $x, x' \in h$, $x^* \prec x$, and $\iota(x^*) = \iota(h)$, then there exists $x^{*\prime} \in H(x^*)$ such that $x^{*\prime} \prec x'$ and $\alpha(y) = \alpha(y')$,

where y and y' are the unique elements of $(P(x) \cup \{x\}) \cap F(x^*)$ and $(P(x') \cup \{x'\}) \cap F(x^{*\prime})$, respectively.

a. In words this means that agents remember everything they did in the past.

b. Translated literally, it says that if x and x' are in the same information set and x^* is a node preceding x at which the same agent chose a certain action, then x' must also be consistent with this choice.

3. Note that $x' \prec x$ is impossible if $x' \in H(x)$, since then we could take $x^* = x'$ in the statement of Perfect Recall and derive an infinite sequence of predecessors of x.

F. The initial assessment is $\rho \in \Delta^\circ(W)$.

1. This probability measure can either be the behavior of nature or the consequence of equilibrium play in a larger game.

2. The assumption that ρ is totally mixed can be dropped without great difficulty for many purposes.

3. Implicitly in this formulation we are assuming that all agents have the same prior beliefs. Clearly this is a natural assumption, but a difference of opinion may easily be a violation of it. It can be dropped without much loss to the formal mathematics, but the resulting concept has not been applied widely. That agents' beliefs are derived from a common prior is known in the literature as consistency in the sense of Harsanyi.

G. An extensive form is a tuple $((T, \prec); A, \alpha; I, \iota; H)$ satisfying the description above.

H. The *utility* or *payoff* is $u = (u_i)_{i \in I}$ where each u_i is a function from Z to $\mathbb{R}$. Thus the space of utilities is $\mathbb{R}^{Z \times I}$.

I. A *game* is a tuple $G = ((T, \prec); A, \alpha; I, \iota; H; \rho; u)$ satisfying the description above.

III. Strategies and Normal Forms

A. The normal form.

1. **Definition:** A *pure strategy* for $i \in I$ is a function $s_i : H^i \to A^i$ such that $s_i(h) \in A(h)$ for all $h \in H^i$

   a. Let S_i be the set of pure strategies for i.

   b. Let $S = \Pi_{i \in I} S_i$ be the set of *pure strategy vectors.*

2. Since $H = \bigcup_{i \in I} H^i$, a pure strategy vector may be thought of as a function $s : H \to A$ with $s(h) \in A(h)$ for all h, and every such function may be thought of as a strategy vector.

Lemma: Given $w \in W$ and $s \in S$, there is a unique $z(w, s)$ such that

$$\alpha(p_{\ell-1}(z(w, s))) = s(H(p_\ell(z(w, s)))), \ell = 1, \ldots, \ell(z(w, s)).$$

Proof: Assumption 2 allows us to choose a sequence $t_0, t_1, \ldots, t_L$ such that $t_0 = w$, and $\alpha(t_\ell) = s(H(t_{\ell-1}))$, $\ell = 1, \ldots, L$, and $t_L \in Z$, and Assumption 1 implies that this sequence is unique. ∎

3. **Definition:** The *normal form* of G is

$$N(G) = (I, (S_i)_{i \in I}, (u_i)_{i \in I})$$

   where $u_i(s) = \Sigma_{w \in W}\, \rho(w) \cdot u_i(z(w, s))$. (Here we are using u_i to denote both the utility associated with a terminal node and the expected utility associated with a strategy vector. The correct interpretation must be inferred from the context.)

B. The reduced normal form.

1. Two pure strategies may be realization equivalent in the same sense as before.
2. One can define equivalence classes and a reduced normal form.
3. Since this will not be important we do not do this formally.

C. The agent normal form.

1. As we noted above, we can identify S and $\Pi_{h \in H} A(h)$.
2. **Definition:** The *agent normal form* of G is

$$AN(G) = (H, \, (A(h))_{h \in H}, \, (u_{\iota(h)})_{h \in H}).$$

   a. Intuitively each information set h is a different agent with the preferences of $\iota(h)$.
3. **Question:** Is the agent normal form an accurate description of the strategic possibilities available to the agents?
   a. For pure strategies it certainly is: each normal form pure strategy can be achieved in the agent normal form and vice versa.
   b. For mixed strategies it is not so clear.
      i. Any vector $(\pi_h \in \Delta(A(h)))_{h \in H^i}$ of mixed strategies for the information sets at which i chooses gives rise to a mixed strategy for i by the formula $\sigma_i(s_i) = \Pi_{h \in H^i} \, \pi_h(s_i(h))$.
      ii. The converse is not true: there are mixed strategies $s_i \in \Delta(S_i)$ that cannot be derived in this way from any agent normal form mixed strategies for the informations sets at which agent i chooses.
   c. Kuhn's Theorem shows that the converse is "almost" true provided the assumption of perfect recall is satisfied.

IV. Kuhn's Theorem

A. **Definition:** A *behavior strategy for agent* i is a vector $\pi_i = (\pi_h)_{h \in H^i}$ of probability distributions $\pi_h \in \Delta(A(h))$. Let Π_i be the set of behavior strategies for i.

1. A *behavior strategy* (for the game) is a vector

$$\pi = (\pi_i)_{i \in I} \in \Pi_i = \Pi_{i \in I} \Pi_i.$$

2. A behavior strategy π_i for i induces a mixed strategy $\sigma_i(\pi_i) \in \Delta(S_i)$ defined by the formula for the probability of the joint occurrence of independent

events:

$$\sigma_i(\pi_i)(s_i) = \Pi_{h \in H^i} \pi_h(s_i(h)).$$

3. The point of Kuhn's Theorem is that an agent does not lose any strategic flexibility by restricting himself to strategies of the form $\sigma_i(\pi_i)$.

B. Probabilities.

1. **Definition:** We say that s_i *allows* t, and write $s_i \to t$, if $s_i(H(p_\ell(t))) = \alpha(p_{\ell-1}(t))$ for all $\ell = 1, \ldots, \ell(t)$ such that $H(p_\ell(t)) \in H^i$. We say that s_i *allows* h, and write $s_i \to h$, if $s_i \to x$ for any $x \in h$.
   a. By Perfect Recall, if $s_i \to h$ then $s_i \to x$ for *all* $x \in h$.
2. If w is an initial node that precedes t, and $\sigma \in \Sigma$, then the *probability of* t *conditional on* w *under* σ is

$$\mathbf{P}^\sigma(t|w) = \Pi_{i \in I} \sigma_i(\{s_i | s_i \to t\}).$$

Kuhn's Theorem: For any $\sigma_i \in \Delta(S_i)$ there is $\pi_i \in \Pi_i$ such that for all $\sigma_{-i} \in \Sigma_{-i}$, all w, and all t with $w \prec t$,

$$\mathbf{P}^{(\sigma_i, \sigma_{-i})}(t|w) = \mathbf{P}^{(\sigma_i(\pi_i), \sigma_{-i})}(t|w).$$

Proof: In view of the formula for $\mathbf{P}^\sigma(t|w)$ it suffices to find π_i with

$$\sigma_i(\pi_i)(\{s_i | s_i \to t\}) = \sigma_i(\{s_i | s_i \to t\})$$

for all nodes t.

We define π_i as follows. If $h \in H_i$ and $\sigma_i(\{s_i | s_i \to h\}) = 0$, then we choose $\pi_h \in \Delta(A(h))$ arbitrarily. If $\sigma_i(\{s_i | s_i \to h\}) > 0$, then π_h is defined by

$$\pi_h(a) = \frac{\sigma_i(\{s_i | s_i \to h \text{ and } s_i(h) = a\})}{\sigma_i(\{s_i | s_i \to h\})} \; (a \in A(h)).$$

In words $\pi_h(a)$ is the probability of choosing action a conditional on allowing h to be reached. This is obviously correct intuitively, but the algebra is rather tedious.

Fix $t \in T$. Let $h_1, \ldots, h_r$ be the information sets in H^i that occur on the way to t, and let a_k be the action chosen at h_k en route, $k = 1, \ldots, r$. We assume that $h_1, \ldots, h_r$ are in order of precedence, so that all s_i allow h_1, and s_i allows h_{k+1} if and only if s_i allows h_k and $s_i(h_k) = a_k$. Moreover,

$$\sigma_i(\pi_i)(\{s_i | s_i \to t\}) = \sigma_i(\pi_i)(\{s_i | s_i(h_k) = a_k,\ k = 1, \ldots, r\}).$$

The events $a_1, \ldots, a_r$ are statistically independent under $\sigma_i(\pi_i)$, so the probability of their joint occurrence is the product of their individual probabilities. Thus

$$\begin{aligned}
\sigma_i(\pi_i)(\{s_i | s_i \to t\}) &= \Pi_{k=1,\ldots,r}\, \pi_{h_k}(a_k) \\
&= \sigma_i(\{s_i | s_i(h_1) = a_1\}) \\
&\quad \cdot \Pi_{k=1,\ldots,r} \frac{\sigma_i(\{s_i | s_i \to h_k \text{ and } s_i(h_k) = a_k\})}{\sigma_i(\{s_i | s_i \to h_k\})} \\
&= \sigma_i(\{s_i | s_i(h_1) = a_1\}) \\
&\quad \cdot \Pi_{k=1,\ldots,r-1} \frac{\sigma_i(\{s_i | s_i \to h_{k+1}\})}{\sigma_i(\{s_i | s_i \to h_k\})} \\
&\quad \cdot \frac{\sigma_i(\{s_i | s_i \to h_r \text{ and } s_i(h_r) = a_r\})}{\sigma_i(\{s_i | s_i \to h_r\})} \\
&= \sigma_i(\{s_i | s_i \to h_r \text{ and } s_i(h_r) = a_r\}) = \sigma_i(\{s_i | s_i \to t\}).
\end{aligned}$$

Of course this calculation is invalid if $\sigma_i(\{s_i | s_i \to h_k\}) = 0$ for some k, but then $\sigma_i(\pi_i)(\{s_i \to t\}) = 0 = \sigma_i(\{s_i | s_i \to t\})$. ∎