Multistage games with observed actions

These games are characterized by:

- The game progresses through a sequence of stages, possibly infinitely many. Often each stage is also a time period, as in repeated games.
- Players move simultaneously within a stage.
- Players know the actions of players in previous stages.

For a finite number of stages, these games can be modeled as extensive form games with imperfect information (imperfect due to the simultaneous moves).

Notation

- $N$: number of players
- $a_i(t)$: action of player $i$ at time $t$ and $a_i(t) \in A_i$
- $h_t = (a_i(s); 1 \leq i \leq N, 0 \leq s \leq t-1)$: history through stage $t-1$.

A (pure strategy) policy for player $i$, $\mu_i$, is a mapping $\mu_i: (t, h_t) \rightarrow A_i$ at stage $t$. history.
Payoff functions:  
\[ J_i(\mu_i, \mu_{-i}) = \sum_{t=1}^{K} u_i(t, \mu_i(t, h_t), \mu_{-i}(t, h_t)) \]

where \( u_i(t, a) \) is the payoff function for stage \( t \).

For a given \( k \) with \( 0 \leq k \leq K \) and history \( h_k \) fixed, the subgame for \((k, h_k)\) is the game with payoff function:

\[ J_i^{(k)}(\mu_i, \mu_{-i} | h_k) = \sum_{t=k}^{K} u_i(t, \mu_i(t, h_t), \mu_{-i}(t, h_t)) \]

(only depends on \( \mu_i(t, h_t) \) for \( t \geq k \))

The strategy profile \((\mu_i)_{i \in S} \cup N)\) is \textit{subgame perfect} if it is a Nash equilibrium for every subgame \((k, h_k)\).
Example

Six rounds of prisoner's dilemma $t = 0, 1, \ldots, 5$

Player 2

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1,1</td>
<td>-1,2</td>
</tr>
<tr>
<td>D</td>
<td>2,-1</td>
<td>0,0</td>
</tr>
</tbody>
</table>

$\text{NE}$

In fact, $D$ is a dominant strategy.

Trigger strategy $\mu_i^T$ for player $i$: Play $C$ if no player has ever played $D$. Otherwise play $D$.

Is $(\mu_1^T, \mu_2^T)$ a NE?

No. In fact, no matter what happens in first 5 rounds, for any NE, both players play $D$ in last round. So also in second to last round.

So for all rounds,

\[ \text{H.E. } \mu_i(t, b_t) = D \text{ for } i \in \{1, 2\}, \text{ and } k \]

is a Nash equilibrium with payoff vector $(0, 0)$.

Payoff vector is $(0, 0)$.
If player i uses 

\[ \mu_i(t, h_t) = d, \]

player i will never encounter a history such that player i himself played c at some time -- even if other player deviates. So \((\mu_1, \mu_2)\) would still be a NE if for one i, \(\mu_i\) were redefined arbitrarily on histories such that player i played c at least one.

For a subgame perfect equilibrium need \((\mu_1, \mu_2)\) to be a NE for every subgame, (even ones that couldn't possibly occur -- think of nodes on tree of extensive form games.) Can see \((\mu_1 = d, \mu_2 = d)\) is the unique subgame perfect equilibrium for six rounds of prisoner's dilemma.
Variation: After each round, roll a die. If 6 appears, stop. Total number of rounds $X \sim B(6, 1/6)$.

Is $(M_i(t, h_t) = d, i = 1, 2)$ still a subgame perfect NE? Yes.

Any others?

How about if both players play trigger strategy? "trigger"

$$J_i(M_i^T, \mu^T) = E[\sum_{t=0}^{X-1} \mathbb{1}_{X=t}]$$

$$= E[X] = 6 = \sum_{t=0}^{\infty} \left(\frac{5}{6}\right)^t$$

Payoff vector is $(6, 6)$.

If other player is known to be using the trigger strategy and you use

\[\underbrace{C, C, C, d, C, C, d, \ldots}_{\text{L times}}\] for an integer $L > 0$.

Start out this way. If you play d once, it's best to play d ever after.

Payoff is

\[1, 1, 1, 2, 0, 0, 0, \ldots\] up until $X-1$

i.e.

$$\sum_{t=0}^{L} \left(\frac{5}{6}\right)^t + 2 \left(\frac{5}{6}\right)^L = \frac{1 - \left(\frac{5}{6}\right)^{L+1}}{1 - \frac{5}{6}} + 2 \left(\frac{5}{6}\right)^L$$

$$= 6 + \left(\frac{5}{6}\right)^L \{2 - 6^2\} < 6.$$
So \((\mu^T, \mu^T)\) is a NE.

It's subgame perfect as well.

True for any continuum probability \(s\)

with \(s > \frac{1}{2}\) (payoff of "cooperate \(k\) times")

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Payoff (vs. (\mu^T_i))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu_i^T)</td>
<td>(\frac{1}{1-s} \leq \frac{1}{2} \leq s &lt; 1)</td>
</tr>
<tr>
<td>(d, d, \ldots)</td>
<td>2</td>
</tr>
<tr>
<td>((\frac{c, c, \ldots, c, d, d, \ldots}{2 \text{ times}}))</td>
<td>(\frac{1}{1-s} + s \left{ 2 - \frac{1}{1-s} \right} )</td>
</tr>
</tbody>
</table>
We can formalize the connection between backward induction and SPE as follows.

**Definition** A strategy profile $\mu^*$ satisfies the one-stage-deviation (OSD) condition if for any $i, k, \tilde{h}_k$ with $0 \leq k \leq K$, if $\bar{\mu}$ agrees with $\mu^*$ except at $(i, \tilde{h}_k)$, then

$$J_i^k(\bar{\mu}, \mu_i^*, \tilde{h}_k^*) \leq J_i^k(\mu_i^*, \mu_i^*, \tilde{h}_k^*). \quad (*)$$

**Note** If $K$ is finite, then $\mu^*$ satisfies OSD condition if and only if $\mu^*$ can be derived using backwards induction.

**Proposition** One-stage-deviation principle ($K < +\infty$) (for multistage games with observed moves)

A strategy profile $\mu^*$ is a subgame perfect equilibrium (SPE) iff $\mu^*$ satisfies the OSD condition.
Proof (only if) If \( \mu^* \) is an SPE then, by definition, for any \( i, k, h_k \) with \( 0 \leq k \leq K \), (*) is true for any \( \tilde{\mu}_i \), including \( \tilde{\mu}_i \) that differs from \( \mu^* \) only at stage \( k \) for history \( h_k \). So \( \mu^* \) satisfies OSD condition.

(i) Suppose \( \mu^* \) satisfies the OSD condition.

Let \( i \) be an arbitrary player and let \( \mu_i \) be an arbitrary strategy for \( i \). To show that \( \mu^* \) is an SPE, it suffices to show that for \( l \leq k \leq K \):

\[
J_i^{(k)} (\mu_i, \mu_i^* | h_k) \leq J_i^{(k)} (\mu_i^*, \mu_i^* | h_k) \quad \text{for all } h_k 
\]

(*)

(**)

We use proof by backwards induction on \( k \).

For the base case, note that (**) is true for \( k = K \) because for each choice of \( h_K \), \( \mu_i \) enters only by its value for stage \( K \) and history \( h_K \), \( \mu_i(K, h_K) \).
Suppose (*) is true for \( k+1 \). Consider an arbitrary choice of \( h_k \) and let \( h_{k+1} \) be the extension of \( h_k \) using \( \mu_i \) and \( \mu_i^* \):

\[
h_{k+1} = (h_k, \mu_i(k, h_k), \mu_i^*(k, h_k)).
\]

Then

\[
J_i^{(k)}(\mu_i, \mu_i^* \mid h_{k+1}) = u_i(k, \mu_i(k, h_k), \mu_i^*(k, h_k))
\]

by induction hypothesis

\[
+ J_i^{(k+1)}(\mu_i, \mu_i^* \mid h_{k+1})
\]

\[
\leq u_i(k, \mu_i(k, h_k), \mu_i^*(k, h_k))
\]

\[
+ J_i^{(k+1)}(\mu_i^*, \mu_i^* \mid h_{k+1})
\]

\[
= J_i^{(k)}(\tilde{\mu}_i, \mu_i^* \mid h_k)
\]

where \( \tilde{\mu}_i \) agrees with \( \mu_i \) at \( (k, h_k) \) and with \( \mu_i^* \) elsewhere. Hence (**) is true for \( k+1 \).

Therefore, by proof by induction, (**) is true for \( 0 \leq k \leq K \), completing the proof.
The OSD principle can be extended to $K = \infty$ under the following condition:

**Definition.** The game is continuous at infinity if

$$\lim_{k \to \infty} \sup_{i} \left| J_i^{(k)} (\mu_h) - J_i^{(k)} (\tilde{\mu} h_k) \right| = 0$$

(True, for example, if stage payoffs have the form $g_i(\mu(t), h_k) s^k$ for discount factor $s$ with $0 < s < 1$.)

**Proposition.** One-stage-deviation principle ($K = \infty$)

If the game is continuous at infinity then $\mu^*$ is an SPE iff it satisfies the OSD condition.

**Proof.** (only if) Same as case $K < \infty$. 

Proof (if) Suppose $\mu^*$ satisfies the OSD condition.

Let $i > 0$. Fix $i$ and a policy $\mu_i$. It suffices to show:

$$J_i^{(k)}(\mu_i, \mu_i^* | h_k) \leq J_i^{(k)}(\mu_i^*, \mu_i^* | h_k) + \varepsilon$$

for all $h_k$ (\ding{55}) for all $i$. By the continuity at infinity condition, there exists $\tilde{k}$ so large that (\ding{55}) is true for all $k > \tilde{k}$.

The backwards induction proof used for $K < \infty$ can be used to show (\ding{55}) for $0 \leq k \leq \tilde{k}$ as well, and hence for all $k$. \hfill \Box
Feasibility theorems (aka folk theorems) for repeated games

Let \( G = (I, (S_i : i \in I), (\mathbf{g}_i : i \in I)) \) be a finite strategic form game and let \( 0 < \delta < 1 \).

The repeated game with stage game \( G \) and discount factor \( \delta \) is the strategic form game with payoff functions

\[
J_i(\mathbf{\mu}) = (1-\delta) \sum_{t=0}^{\infty} \delta^t g_i(\mathbf{\mu}(t), h(t))
\]

Here \( \mathbf{\mu} \) is a strategy profile

\[
\mathbf{\mu} = (\mu_i(t, h_t): i \in I, t \geq 0)
\]

Note that \( (1-\delta) \sum_{t=0}^{\infty} \delta^t = 1 \), so if each player uses a constant strategy \( \mu_i(t) = \alpha_i \)

then \( J_i(\mathbf{\mu}) = g_i(\mathbf{\mu}) \) for all \( i \).
F2

From example of prisoner's dilemma, we've seen existence of multiple equilibria.
Feasibility theorems address what payoff vectors can be realized by the repeated game for \( S \) sufficiently close to one.

Let \( V = (V_i; i \in I) \) be defined by

\[
V_i = \max_{x_i} \min_{x_{-i}} g_i(x_i, x_{-i})
\]

= \min_{x_{-i}} \max_{x_i} \text{ value for player } i \in I \in \mathcal{G}.

By repeatedly playing \( \hat{x}_i \in \arg\max_{x_i} \min_{x_{-i}} g_i(x_i, x_{-i}) \), player \( i \) can insure \( J_i(\mu) \geq V_i \), no matter what other players play. So

\[
J(\mu^{NE}) \geq V \quad \text{for any } NE \mu^{NE}.
\]
F3

- A vector \( v \) is \textit{individually rational (IR)} if \( v_i \geq v_i \) for \( i \in I \), and \textit{strictly IR} if \( v_i > v_i \) for all \( i \).

- A vector \( v \) is a \textit{feasible} payoff vector for game \( G \) if

\[
v \in \text{convex hull} \left\{ g(x) : x \in X_S, i \in I \right\}
\]

Feasible payoffs are those that can be achieved for \( G \) using mixed strategies and randomization based on publicly available randomness.

\[\begin{array}{c|cc}
 & d & c \\
\hline
\text{d} & 1,1 & -1,2 \\
\text{c} & 2,-1 & 0,0 \\
\end{array}\]

\( V = (10,10) \) \( \leftarrow \) (Happens to be payoff for unique NE)
Theorem (Nash) If \( \bar{V} \) is feasible and strictly IR then there exists \( \bar{\delta} \in (0, 1) \) so for any \( \delta \in (\bar{\delta}, 1) \), there exists an NE profile \( \mu \) so that \( \bar{V} = J(\mu) \).

Proof (Assuming public randomness available -- can be simulated for \( \delta \) large enough)

There is a random strategy profile \( \alpha \) for \( G \) so that \( \bar{V} = E[J(\alpha)] \). Use trigger strategies following \( \alpha \), which switch to punishing first player to violate this script.

Objective -- the NE is typically not subgame perfect. Punishing one player could be costly to the others. (Not for prisoners dilemma)
Theorem (Friedman) Let $v^{NE}$ be a payoff vector for some NE $x^{NE}$ of the stage game $G$. If $v$ is a feasible vector such that $v_i > v_i^{NE}$ for all $i$, then there exist $\overline{s} \in (0, 1)$, so that for any $\bar{s} \in (\overline{s}, 1)$, there exists a subgame perfect profile $\mu$ with $v = J(\mu)$. 

Proof (Again, under publicly observable randomness) Player $i$ uses trigger strategy starting with $x_i$, switching to $x_i^{NE}$ forever if one player deviates from using $v_i$ ($x$ is random profile for $G$ with $v = E[g(x)]$.)
Theorem: Under same conditions as

(Fudenberg, Maskin 86)
Nash theorem, if the dimension of the
set of feasible vectors is equal to the
number of players, then conclusion of
Nash theorem holds with ρ being
subgame perfect. (See F+T for proof.)

Idea is to incentivize other players
to punish first player to deviate from
script for a trigger policy. Easier
to prove under availability of public
randomness, but still complicated.
Bayesian games (Bayes-Nash equilibrium for games with incomplete information)

Def A Bayesian game is given by

\[ G = (I, (S_i)_{i \in I}, (\Theta_i)_{i \in I}, U_i, s, (\Theta)) \]

- Space of types for player \( i \)

Marginals: \( p(\Theta_i) = \sum_{\Theta_i} p(\Theta_i, \Theta); \) Assume \( p(\Theta_i) > 0 \)

so \( p(\Theta_i | \Theta_i) \) is well defined. A pure strategy for a player \( i \) is a mapping \( s_i: \Theta_i \rightarrow S_i \).

A strategy profile \( (s_i)_{i \in I} \) is a Bayesian (or Bayes-Nash) equilibrium if for each \( i \) and \( \Theta_i \in \Theta_i \),

\[ s_i(\Theta_i) \in \arg \max_{a_i' \in S_i} \max_{\Theta_i'} \sum_{\Theta_i} p(\Theta_i, 1 \Theta_i) U_i(a_i', s_i(\Theta_i), \Theta_i, \Theta_i') \]

In applications the type \( \Theta_i \) of player \( i \) is viewed as private information of player \( i \), determined before the game began.

Harsanyi (73) suggests viewing the situation
as a two stage game in which nature selects the vector of types \((E_i)_{i \in I}\) at the players' stage 1 using the known joint distribution \(\mathbf{p}\), each player learns his own type, and the players play the game with action spaces \((\mathcal{S}_i)_{i \in I}\) and payoff functions \(u_i(s, E)\). Player \(i\) doesn't know \(\Theta_i\), but knows its conditional distribution \(p(E_i \mid \Theta_i) = \frac{p(E_i, E)}{\sum_{E'} p(E_i, E')}\).

If the game following assignment of \(\Theta_i\)'s by nature has multiple stages, strategies could involve learning about the types of others.

Bayes equilibrium exists (special case of result for extensive form games or use fixed point proof.)
The half-Kuhn homework problem could be viewed as a Bayesian game. It began by dealing one card from a deck of three cards (\{L, M, H\}) to each of player 1 and player 2. The card of player 1 can be viewed as a type. In that example, types are dependent.

Example: A Cournot game with incomplete information.
(see problem set 5) Given \( C_1 \approx 79,700, \ 0 < p < 1. \)

Suppose production cost (per unit) for player 1 is \( C_i \) with probability \( p \) \((G_i = 1)\)
\[ \begin{align*}
0 & \quad \text{w} \quad \text{p} \quad \text{f} \quad \text{r} \\
1 - p & \quad \text{f} \quad \text{r} \quad \text{p} \quad \text{w}
\end{align*} \]

and for player 2 is \( C \) (with probability one).
Player 1 strategy is a pair \((\bar{b}_1, \bar{b}_2)\)

Let \(a \in \mathbb{R}\) be a number \(\bar{b}_2\)

\[
\begin{align*}
\bar{b}_1, &= \text{arg max}_{\tilde{b}_1} \tilde{b}_1 (a - \tilde{b}_2 - \tilde{b}_1) \\
&= \left( \frac{a - \bar{b}_2}{2} \right) + \\
\bar{b}_1, &= \text{arg max}_{\tilde{b}_1} \tilde{b}_1 (a - \tilde{b}_2 - \tilde{b}_1) \\
&= \left( \frac{a - \bar{b}_2}{2} \right) + \text{ (player 1 produces less if his cost is larger)} \\
\bar{b}_2 &= \text{arg max}_{\tilde{b}_2} (1-p) \tilde{b}_2' (a - \bar{b}_1, -c - \tilde{b}_2') \\
&= \left( \frac{a - \bar{b}_1}{2} \right) + p \tilde{b}_2' (a - \bar{b}_1, -c - \tilde{b}_2') \\
&= \text{arg max}_{\tilde{b}_2} (1-p) \tilde{b}_2' (a - \bar{b}_1, -c - \tilde{b}_2') = \left( \frac{a - \bar{b}_1}{2} \right)
\end{align*}
\]

where \(\bar{b}_1 = (1-p) \bar{b}_{1,0} + p \bar{b}_{1,1}\). Note: \(\bar{b}_1 = \frac{a-pc-\bar{b}_2}{2}\)

if "( )" above are inactive. So we have fixed point equations for \((\bar{b}_1, \bar{b}_2)\). ... continued in ps 5 # 4.
Fudenberg and Tirole introduced a notion of perfect Bayes equilibrium (see their book). Their PBE is a somewhat weaker property than sequential equilibrium in the equivalent extended game (with complete information, using first node controlled by nature.)

(Discussed in Myerson book.)
The Vickrey-Clarke-Groves (UCG) mechanism.

Let $C$ represent a set of possibilities for an allocation or social choice, affecting $N$ players.

Examples:

(a) $C$ could be a set of possible locations of a new school in a community.

(b) Auction of a single object.

$$C = \{1, \ldots, N\}; \text{ i means player } i \text{ gets object.}$$

(c) Auction of objects in a finite set $O$.

$$C = \{(A_1, \ldots, A_n); A_i \in O, A_iA_j = \emptyset\}$$

Suppose player $i$ has a utility function $U_i(x)$. A mechanism $M = (g, m_1, \ldots, m_N)$ takes a vector of bids $(s_1, \ldots, s_N)$ of the players and assigns an allocation $g(s) \in C$ and a set of payments $(m_1(s), \ldots, m_N(s))$ to be made by the players.
A mechanism induces a game among the players. The payoff function for player \( i \) is

\[ \pi_i(s) = u_i(g(s)) - m_i(s) \]

If \( x^* = g(s^{NE}) \), where \( s^{NE} \) is a Nash equilibrium for the game, and \( x^* \) satisfies some property \( P \), we say that the mechanism is a Nash implementation of property \( P \).

The UCG mechanism described next gives an implementation in weakly dominated strategies (so stronger than Nash equilibrium) of maximum welfare allocation.
The (social) welfare of an assignment $x$ is defined by $W(x) = \sum_{i=1}^{n} U_i(x)$.

The set of maximum welfare assignment: $\arg\max_{x \in \mathcal{X}} W(x)$

**VCG mechanism**

Space of bids is space of utility functions. So use $\hat{U}_i$ instead of $S_i$ for notation.

Given the bids $\hat{U}_i$, $x^*$ is selected with $x^* \in \arg\max_{x \in \mathcal{X}} \sum_{i=1}^{n} \hat{U}_i(x)$ (i.e. $x^*$ maximizes welfare for repeated utility functions).

Payments

$$m_i(\hat{U}) = -\sum_{j:j \neq i} \hat{U}_j(x^*) + t_i(\hat{U}_{-i})$$

Claim: Bidding $U_i$ (i.e. truthful bidding) is a weakly dominant strategy of player $i$. 
Notes: (a) \( m_i(\hat{u}) \) depends on \( \hat{u}_i \) only through the fact \( \hat{u}_i \) influences the selection of \( x^* \).

(b) Above, \( \hat{u}_{-i} \) stands for the vector of functions \( \hat{u}_{-i} = (\hat{u}_j : j \neq i) \).

It is not the vector of functions evaluated at \( x^* \). So \( \hat{u}_{-i} \) does not depend on the bid, \( \hat{u}_i \), of player \( i \).

A common choice for \( t_i \) is discussed below.
Proof of claim: The payoff of player $i$ is given by

$$\Pi_i(\hat{x}) = \Pi_i(x^*) + \sum_{j \neq i} \hat{\Pi}_j(x^*) - t_i(\hat{\hat{u}}_i)$$

social welfare of $x^*$ not influenced by $u_i$

and reports of other players using $u_i$ and reports of other players

Thus, the payoff of player $i$ is the social welfare (computed using $u_i$ and $\hat{\hat{u}}_i$) minus a term not depending on $\hat{\hat{u}}_i$. No matter what $\hat{\hat{u}}_i$ was reported by other players, by reporting $u_i$ truthfully (i.e. taking $\hat{\hat{u}}_i = u_i$) player $i$ directs the mechanism to select $x^*$ to maximize $\Pi_i(\hat{x})$ with respect to $\hat{\hat{u}}_i$.

(Basically, the payment rule aligns the interests of each player $i$ with the social good.)
A common choice for $t_i(\hat{\mathbf{u}})$ is to choose $\hat{\mathbf{u}}_i$ so as to maximize social welfare for the other players, based on their reported bids:

$$\max_{\mathbf{x} \in \Xi} \sum_{j \neq i} \hat{\mathbf{u}}_j(\mathbf{x})$$

This is a function of $\hat{\mathbf{u}}_{-i}$, and is a common choice for $t_i(\hat{\mathbf{u}}_{-i})$. For this choice,

$$(*) \quad m_i(\hat{\mathbf{u}}) = \left( \max_{\mathbf{x} \in \Xi} \sum_{j \neq i} \hat{\mathbf{u}}_j(\mathbf{x}) \right) - \sum_{j \neq i} \hat{\mathbf{u}}_j(\mathbf{x}^*)$$

Then $m_i(\hat{\mathbf{u}}) \geq 0$ and it can be viewed as the loss in the total welfare of the other players caused by the actions/existence of player $i$. 

Recovering second price auction from VCG

Recall the example of auctioning a simple object: \( C = \{1, \ldots, N\} \). Suppose player \( i \) has a value \( V_i \) for the object. Then \( u_i(x) = V_i I_{\{x = i\}} \). Each player \( i \) reports a value \( \hat{V}_i \). Then

\[
\hat{x}^* = \text{arg max}_{1 \leq x \leq N} \hat{V}_x
\]

For the choice (\( \ast \)), the payments are

\[
m_i(\hat{x}) = \begin{cases} 
\max_{j: j \neq i} \hat{V}_j & \text{if } \hat{x}^* = i \\
0 & \text{else}
\end{cases}
\]

That is, the payment of the winner is the highest bid of the other players. Losers make zero payment. The winner is a player with a maximum bid. Thus, VCG reduces to the Vickrey second price auction.