I. Introduction
A. We now begin the study of solution concepts for games in normal form.
1. We begin with the weakest concepts and progress to more and more restrictive solution concepts.
2. This way of ordering the material is natural in the absence of a consensus about the “correct” solution concept.
3. It is also useful from the point of view of applications.
   a. Powerful (i.e. restrictive) solution concepts allow one to say things about games that would otherwise be beyond analysis.
   b. On the other hand, when one can derive conclusions from weak solution concepts, the conclusions are more compelling insofar as they do not depend on strong assumptions about behavior.
B. Today we are concerned with solution concepts that presuppose that the structure of the game is common knowledge, but do not assume that the agents’ strategies are common knowledge. (This is a kind of “rational expectations” assumption.)
1. The simplest such notion is strict dominance.
   a. If there are two strategies for an agent such that one has a higher payoff than the other no matter what the other agents do, then the better strategy strictly dominates the worse strategy.
   b. We can eliminate all dominated strategies, then all strategies that are dominated in the truncated game, and so on.
c. A closely related concept is *strict mixture domination* in which a pure strategy is strictly dominated by a mixed strategy (a probability distribution on pure strategies).

2. When one strategy for an agent gives at least as high a payoff as another, regardless of what the other agents do, and in some circumstances gives a strictly higher payoff, then we say that the first strategy *weakly dominates* the second.

   a. This concept is less well behaved than strict dominance.

3. It turns out that the restrictions on behavior imposed by common knowledge rationality are somewhat more restrictive than iterative elimination of strictly dominated strategies.

   a. The details have been worked out by David Pearce and Douglas Bernheim.

   b. Their solution concept is known as *rationalizability*.

II. The Notation for Normal Form Games

   A. The set of *agents* is $I = \{1, \ldots, n\}$—this will be fixed throughout the course.

   B. An *n-person normal form game* is a 2n-tuple

\[ N = (S_1, \ldots, S_n; u_1, \ldots, u_n). \]

1. Each $S_i$ is a nonempty finite set of *pure strategies* for agent $i$.

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

3. $u_i: S \to \mathbb{R}$ is agent $i$’s *utility* or *payoff function*.

   C. We adopt the following notation for “everyone but $i$.”

1. $S_{-i} = S_1 \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_n$.

2. If $s_i \in S_i$ and $s_{-i} \in S_{-i}$, then $(s_i, s_{-i})$ is the obvious strategy vector.
III. Strict Dominance

A. We say that $s_i$ strictly dominates $t_i$, and that $t_i$ is strictly dominated by $s_i$, if $u_i(s_i, r_{-i}) > u_i(t_i, r_{-i})$ for all $r_{-i} \in S_{-i}$.

B. We say that $s_i$ is strictly dominant if it strictly dominates all other strategies in $s_i$.

   1. The strongest possible solution concept, dominant strategy equilibrium, is that everyone plays a strictly dominant strategy.

   2. Though seemingly unlikely, this is not unheard of in modelling praxis.

   3. The following example is the famous Prisoner’s Dilemma.

      \[
      \begin{array}{ccc}
      & C_2 & D_2 \\
      C_1 & (2, 2) & (0, 3) \\
      D_1 & (3, 0) & (1, 1) \\
      \end{array}
      \]

      a. The following story is told to motivate this game. The police apprehend two people who they suspect of committing a crime together. They are interrogated in separate rooms, and each is told that if both turn state’s evidence they will receive moderately severe sentences, if neither turns state’s evidence they will receive mild sentences, and if one turns state’s evidence while the other does not then the one who turns state’s evidence will go free and the other one will receive a very severe sentence.

      b. For each individual it is a dominant strategy (regardless of whether they are guilty!) to “defect,” i.e. turn state’s evidence.

      c. Both defecting is a Pareto inferior outcome since they are both better off if they both “cooperate,” i.e. neither turns state’s evidence.

C. If one agent has a dominated strategy then other agents will not expect him to play it, and it may be the case that some strategies of the other agents become dominated when this strategy is eliminated.
1. In the following example iterative elimination of dominated strategies determines a unique outcome.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>(2, 2)</td>
<td>(2, 0)</td>
<td>(4, -2)</td>
</tr>
<tr>
<td>B</td>
<td>(1, 1)</td>
<td>(3, 0)</td>
<td>(6, -2)</td>
</tr>
<tr>
<td>C</td>
<td>(-1, 7)</td>
<td>(-1, 8)</td>
<td>(5, -2)</td>
</tr>
</tbody>
</table>

2. It is important that iterative elimination of strictly dominated strategies is a well defined procedure in the sense that the end result does not depend on the order of elimination.

**Lemma:** Suppose that $r_{j_1}^1, \ldots, r_{j_{\ell}}^\ell$ is a sequence of pure strategies that can be iteratively eliminated by strict dominance. Then, for each $h = 1, \ldots, \ell$, there is some

$$q_{j_h} \in S_{j_h} \setminus \{r_{j_1}^1, \ldots, r_{j_{\ell}}^\ell\}$$

such that $u_{j_h}(q_{j_h}, t_{-j_h}) > u_{j_h}(r_{j_h}^h, t_{-j_h})$ for all $t_{-j_h} \in S_{-j_h}$ that have no components in the list $r_{j_1}^1, \ldots, r_{j_{\ell}}^\ell$.

**Proof:** Clearly this is true when $\ell = h$: $r_{j_{\ell}}^\ell$ was eliminated because it was strictly dominated. By induction on $\ell$ we may assume that for each of $h = g + 1, \ldots, \ell$, there is some

$$q_{j_h} \in S_{j_h} \setminus \{r_{j_1}^1, \ldots, r_{j_{\ell}}^\ell\}$$

such that $u_{j_h}(t_{-j_h}, q_{j_h}) > u_{j_h}(t_{-j_h}, r_{j_h}^h)$ for all $t_{-j_h} \in S_{-j_h}$ that have no components in the list $r_{j_1}^1, \ldots, r_{j_{\ell}}^\ell$. We know that $r_{j_g}^g$ is strictly dominated in the game resulting from eliminating $r_{j_1}^1, \ldots, r_{j_{g-1}}^{g-1}$. If it is dominated in a strategy that is not in the list $r_{j_{g+1}}^{g+1}, \ldots, r_{j_{\ell}}^\ell$, then we may take $q_{j_g}$ to be this strategy. Otherwise the strategy dominating $r_{j_g}^g$ is $r_{j_h}^h$ for some $h = g + 1, \ldots, \ell$. In this case we may take $q_{j_g}$ to be $q_{j_h}$, since $j_h = j_g$ and

$$u_{j_g}(q_{j_g}, t_{-j_g}) > u_{j_g}(r_{j_h}^h, t_{-j_g}) > u_{j_h}(r_{j_g}^g, t_{-j_g})$$

for all $t_{-j_g} \in S_{-j_g}$ that have no components in the list $r_{j_1}^1, \ldots, r_{j_{\ell}}^\ell$. ■
Theorem: Suppose $r^1_{j_1}, \ldots, r^\ell_{j_\ell}$ and $s^1_{k_1}, \ldots, s^m_{k_m}$ are two sequences of pure strategies that can be iteratively eliminated by strict dominance, then $r^1_{j_1}, \ldots, r^\ell_{j_\ell}, s^1_{k_1}, \ldots, s^m_{k_m}$ is also such a sequence. (Here it may happen that various $s^h_{k_h}$ have already been eliminated as members of the sequence $r^1_{j_1}, \ldots, r^\ell_{j_\ell}$.)

Proof: Arguing by induction on $k$, it suffices to show that $s^m_{k_m}$ is dominated in the game obtained by eliminating $r^1_{j_1}, \ldots, r^\ell_{j_\ell}, s^1_{k_1}, \ldots, s^{m-1}_{k_{m-1}}$. The only way that this can possibly fail to be the case is if all the pure strategies that dominate $s^m_{k_m}$ in the game obtained by eliminating $s^1_{k_1}, \ldots, s^{m-1}_{k_{m-1}}$ are in the list $r^1_{j_1}, \ldots, r^\ell_{j_\ell}$. But for any one of them, say $r^q_{k_m}$, the Lemma implies the existence of a $q_{k_m} \in S_{k_m} \setminus \{r^1_{j_1}, \ldots, r^\ell_{j_\ell}, s^1_{k_1}, \ldots, s^{m-1}_{k_{m-1}}\}$ such that $u_{k_m}(t_{-k_m}, q_{k_m}) > u_{k_m}(t_{-k_m}, r^q_{k_m})$ for all $t_{-k_m} \in S_{-k_m}$ that have no components in the list $r^1_{j_1}, \ldots, r^\ell_{j_\ell}, s^1_{k_1}, \ldots, s^{m-1}_{k_{m-1}}$. Since $r^q_{k_m}$ strictly dominates $s^m_{k_m}$ in the game obtained by eliminating $s^1_{k_1}, \ldots, s^{m-1}_{k_{m-1}}$, it follows that $q_{k_m}$ strictly dominates $s^m_{k_m}$ in the game obtained by eliminating $r^1_{j_1}, \ldots, r^\ell_{j_\ell}, s^1_{k_1}, \ldots, s^{m-1}_{k_{m-1}}$. □

IV. Mixed Strategies and Expected Payoffs

A. A mixed strategy for an agent is a probability distribution over his pure strategies.

1. Typically we denote a mixed strategy for $i$ by $\sigma_i \in \Delta(S_i)$.
2. $\Sigma = \Delta(S_1) \times \cdots \times \Delta(S_n)$ is the space of mixed strategy vectors.
3. Let $\Sigma_{-i} = \Delta(S_1) \times \cdots \times \Delta(S_{i-1}) \times \Delta(S_{i+1}) \times \cdots \times \Delta(S_n)$.
4. If $\sigma_i \in \Delta(S_i)$ and $\sigma_{-i} \in \Sigma_{-i}$, then $(\sigma_i, \sigma_{-i})$ is the obvious mixed strategy vector.
5. The expected payoff for agent $i$ at a mixed strategy vector $\sigma$ is $u_i(\sigma) = \sum_{s \in S} \left( \prod_{j \in I} \sigma_j(s_j) \right) \cdot u_i(s) = \sum_{s_i \in S_i} \sigma_i(s_i) \cdot u_i(s_i, \sigma_{-i})$. 

5
B. We say that \( \tau_i \) is a best response for \( i \) to \( \sigma \in \Sigma \) (actually to \( \sigma_{-i} \in S_{-i} \)) if
\[
u_i(\tau_i, \sigma_{-i}) \geq \nu_i(p_i, \sigma_{-i}) \quad \text{for all} \quad p_i \in \Delta(S_i).
\]
1. Note that \( \tau_i \) is a best response to \( \sigma \) if and only if \( \tau_i(s_i) > 0 \) implies
\[
u_i(s_i, \sigma_{-1}) \geq \nu_i(r_i, \sigma_{-1}) \quad \text{for all} \quad r_i \in S_i.
\]
2. For \( \sigma \in \Sigma \) we let \( BR_i(\sigma) \) be the set of agent \( i \)'s best responses to \( \sigma \).
3. For \( \sigma \in \Sigma \) we let \( BR(\sigma) = \prod_{i \in I} BR_i(\sigma) \). The correspondence
\[
BR: \Sigma \to \Sigma
\]
is called the best response correspondence.

C. Mixture domination.

1. In some cases a pure strategy is strictly dominated by a mixed strategy.

\[
\begin{array}{ccc}
2 & a & b \\
1 & & \\
A & (1, 1) & (1, 6) \\
B & (0, 3) & (3, 4) \\
C & (3, 5) & (0, 2)
\end{array}
\]

2. Iterative elimination of strictly mixture dominated strategies is a well defined procedure in the sense of all complete sequences of eliminations having the same final game, for essentially the reasons given above.

a. Exercise: Suppose \( s_i \) is mixture dominated, say by \( \sigma_i \in \Delta(S_i) \), \( t_i \) is mixture dominated, say by \( \tau_i \in \Delta(S_i) \), and \( \sigma_i(t_i) > 0 \). Construct a mixed strategy \( \rho_i \) with \( \rho_i(t_i) = 0 \) that mixture dominates \( s_i \).

3. Mixture domination involves a stronger behavioral hypothesis than strict domination by pure strategies insofar as it depends on the cardinal nature of utility. Domination by pure strategies depends only on the ordinal preferences over pure outcomes.
V. Weak Domination

A. We say that \( s_i \) weakly dominates \( t_i \), and \( t_i \) is weakly dominated by \( s_i \), if 
\[ u_i(s_i, r_{-i}) \geq u_i(t_i, r_{-i}) \]
for all \( r_{-i} \in S_{-i} \) with strict inequality for some \( r_{-i} \).

B. One might think that rationality should preclude using weakly dominated strategies (why take chances?) but the issue is not clear cut and is somewhat controversial among game theorists.

1. Different sequences of eliminations can result in different final games.

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>a</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>(2,1)</td>
<td>(1,0)</td>
<td>(0,0)</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>(2,1)</td>
<td>(0,0)</td>
<td>(1,0)</td>
<td></td>
</tr>
</tbody>
</table>

2. If \( s_i \) weakly dominates \( t_i \), but those \( r_{-i} \) with \( u_i(s_i, r_{-i}) > u_i(t_i, r_{-i}) \) are thought to be precluded by rationality, then strictly speaking agent \( i \) should be indifferent between \( s_i \) and \( t_i \).

VI. Rationalizability

A. In the traditional interpretation of the mathematical objects called games, the interaction occurs only once, and the agents have no basis for forming expectations about the behavior of others aside from the common knowledge of rationality. (Another possible interpretation will be advanced shortly.)

B. This leads to the following procedure for solving a game worked out independently by Douglas Bernheim and David Pearce.

1. Let \( H_{i0} = \Delta(S_i) \), \( i = 1, \ldots, n \).
2. Let \( J_{i1} = \{ \sigma_i | \sigma_i \in BR_i(\tau) \text{ for some } \tau \in \prod_{j \in I} H_{j0} \} \), \( i = 1, \ldots, n \).
3. Let \( H_{i1} = \text{con}(J_{i1}) \), \( i = 1, \ldots, n \).
a. In general, if $X$ is a subset of $\mathbb{R}^\ell$, then the convex hull of $X$ is

$$
\text{con}(X) = \left\{ \sum_{j=1,\ldots,r} \alpha_j x_j \mid r \text{ is a natural number, } x_1, \ldots, x_r \in X, \alpha_1, \ldots, \alpha_r \in [0,1], \text{ and } \sum_{j=1,\ldots,r} \alpha_j = 1 \right\}.
$$

b. **Exercise:** Prove that $\text{con}(X)$ is the smallest convex set containing $X$. (Hint: first prove that intersections of convex sets are convex.)

c. Here the idea is that, although only strategies in $J_{i1}$ are rational for $i$, other agents’ beliefs about agent $i$’s behavior may be elements of $H_{i1} - J_{i1}$.

4. If $H_{1k-1}, \ldots, H_{nk-1}$ have been defined, let

$$
J_{ik} = \left\{ \sigma_i \mid \sigma_i \in BR(\tau) \text{ for some } \tau \in \prod_{j \in i} H_{jk-1} \right\}
$$

and let $H_{jk} = \text{con}(J_{ik}), i = 1, \ldots, n$.

5. **Theorem:** (Bernheim and Pearce) For each $i$ and $k$, $H_{ik} = \Delta(T_i)$ for some nonempty $T_i \subset S_i$. There is an integer $K$ such that $H_{ik} = H_{iK}$ for all $i$ and $k \geq K$. Moreover, the sets $H_{iK}$ “do not depend on the order of elimination.”

6. The sets $J_{iK}$ are called the sets of *rationalizable strategies*.

7. In two player games rationalizability is equivalent to iterative elimination of strictly mixture dominated strategies (this is a consequence of the separating hyperplane theorem), but this is not true if there are more than two players. In the following example due to Pearce there are three agents with strategy sets $S_1 = \{\ell, m, r\}$, $S_2 = \{U, D\}$, and $S_3 = \{H, T\}$. The following matrices give only the payoffs for agent
1 since the other agents’ payoffs are irrelevant.

\[
\begin{array}{c|cc}
 & H & T \\
\hline
U & 6 & 6 \\
D & 10 & 0 \\
r & 0 & 10 \\
m & 10 & 10 \\
\ell & 6 & 6 \\
\end{array}
\]

a. Here \( \ell \) is not dominated, nor is it strictly mixture dominated since any mixture of \( m \) and \( r \) has an expected payoff no greater than 5 in either the case \((H, U)\) or the case \((T, D)\).

b. Nonetheless \( \ell \) is not rationalizable, since if it were rationalizable there would be \( p = \text{Prob}(U) \) and \( q = \text{Prob}(H) \) such that \( 6 \geq 10 - \max\{1 - pq, 1 - (1 - p)(1 - q)\} \), but we can show that \( \max\{1 - pq, 1 - (1 - p)(1 - q)\} \geq \frac{3}{4} \) for all \( 0 \leq p, q \leq 1 \). If \( p + q \leq 1 \) then \( pq \leq p(1 - p) = \left( \frac{1}{2} - \left( \frac{1}{2} - p \right) \right) \cdot \left( \frac{1}{2} + \left( \frac{1}{2} - p \right) \right) = \frac{1}{4} - \left( \frac{1}{2} - p \right)^2 \leq \frac{1}{4} \), so that \( 1 - pq \geq \frac{3}{4} \), and if \( p + q \geq 1 \) then \( (1 - p) + (1 - q) \leq 1 \) and \( 1 - (1 - p)(1 - q) \geq \frac{3}{4} \).