Handout on Rationalizability[1]

Deleting strategies that are never a best response

Before we use best responses to find the Nash equilibrium of a game, we can apply this concept using a similar approach as that of strictly dominated strategies in the last chapter. Specifically, when applying IDSDS, we argued that, if a player finds one of his strategies to be strictly dominated, we can essentially delete it from the strategy set of this player.
We can now apply a similar approach, arguing that, if a player never uses one or more of his available strategies as a best response to his opponents' strategies, we can label these strategies as "never a best response" (NBR) and delete them from his strategy set. We formally define this concept below.

Never a best response (NBR). Strategy s_i is never a best response if

$$u_i(s_i, s_{-i}) > u_i(s_i', s_{-i}) \text{ for every } s_i' \neq s_i$$

does not hold for any strategy profile of his rivals, s_{-i}.

Alternatively, if a strategy s_i is NBR, there are no beliefs that player i can sustain about how his opponents behave that would lead him to use that strategy. In other words, player i cannot rationalize (explain) why would he ever choose strategy s_i since it is never a best response to his opponents' choices. As a result, we can delete s_i from player i's strategy set S_i.
It is easy to prove that, if player i finds that strategy s_i is strictly dominated by $s_i' \neq s_i$, then s_i yields a strictly lower payoff than s_i' regardless of the specific strategy profile that his opponents select. As a consequence, strategy s_i cannot be a best response against any strategy profile of i's opponents, that is, s_i is NBR.
We can go through an iterative process —analogous to IDSDS— but identifying strategies that are NBR for either player rather than strategies that are strictly dominated, as we list below. This iterative process is known as "rationalizability" because it finds strategies that player i can rationalize (explain) to be best responses to at least one of his opponents' strategies.

Tool 3.2. Applying rationalizability:

Step 1 Starting with player i, delete any strategies that are NBR for him, obtaining the reduced strategy set S_i', where $S_i' \subset S_i$. (This step only uses the definition of rationality as no player would choose a strategy that is NBR.)

Step 2 Using common knowledge of rationality, we can continue the above reasoning, arguing that player $j \neq i$ can anticipate player i's best responses and, as a consequence, the strategies that are NBR for player i, deleting them from S_i.

[1]Felix Munoz-Garcia, School of Economic Sciences, Washingon State University, 103H Hulbert Hall, Pullman, WA 99164-6210, fmunoz@wsu.edu.

Given this reduced strategy space $S_i' \subset S_i$, player j can now examine his own best responses to player i, seeking to identify if one or more are never used, and further restricting her strategy space to $S_j' \subset S_j$.

Step 3 We obtain the Cartesian product $S_i' \times S_j'$, representing the remaining rationalizable strategies after deleting NBRs for two steps. Player i then finds if some of his strategies in this reduced game are NBR, deleting them from her strategy set S_i', obtaining $S_i'' \subset S_i'$

Step k The process continues until we cannot find more strategies that are NBR for either player.

The strategy profile (or set of strategy profiles) that survives this iterative process are referred as "rationalizable" strategy profiles because every player can sustain beliefs about his rival's behavior (i.e., which strategies his rivals choose) that would lead him to best respond using one of these surviving strategies.

Example 3.1 below applies rationalizability to the same payoff matrix as example 2.2, showing that both IDSDS and rationalizability produce the same equilibrium outcomes. It is easy to show that, in games with two players, both solution concepts yield identical equilibrium results. However, in games with three or more players, equilibrium outcomes do not necessarily coincide, with rationalizability producing more precise equilibrium outcomes than IDSDS. That is, for a given strategy profile s,

$$s \text{ is rationalizable } \underset{\not\Leftarrow}{\Longrightarrow} s \text{ survives IDSDS}$$

For more details, see Pearce (1984) and for an example, see Fudenberg and Tirole (1995, pages 51-53 and 62-63).

Example 3.1. Rationalizability and IDSDS. Consider the payoff matrix in example 2.2, which we reproduce in matrix 3.2a for easier reference.

		Firm 2	
		h	l
Firm 1	H	$4, 4$	$0, 2$
	M	$1, 4$	$2, 0$
	L	$0, 2$	$0, 0$

Matrix 3.2a. Applying rationalizability - First step.

Starting with firm 1, it is easy to see that row L is NBR: when firm 2 chooses h, firm 1's best response is H; while when firm 2 chooses l, firm 1's best response is M. In other words, firm 1 does not have incentives to respond with L regardless of the believes this firm sustains on firm 2's behavior. After deleting row L from firm 1's strategy space, S_1, we obtain a reduced strategy space $S_1' = \{H, M\}$, as depicted in matrix 3.2b.

		Firm 2	
		h	l
Firm 1	H	$4, 4$	$0, 2$
	M	$1, 4$	$2, 0$

Matrix 3.2b. Applying rationalizability - Second step.

We can now examine firm 2's best responses: when firm 1 chooses H (top row), firm 2's best response is h; and, similarly, when firm 1 chooses M (bottom row), firm 2's best response is h. Therefore, we can claim that strategy l is NBR for firm 2, entailing a reduced strategy space $S_2' = \{h\}$. As a result, the Cartesian product $S_1' \times S_2'$ has only two remaining strategy profiles, as illustrated in matrix 3.2c.

		Firm 2
		h
Firm 1	H	$4, 4$
	M	$1, 4$

Matrix 3.2c. Applying rationalizability - Third step.

Firm 2 only has one available strategy at this point, $s_2 = h$, implying that firm 1's best response is H and, therefore, strategy M is NBR at this stage of our analysis. Hence, firm 1's strategy set reduces to $S_1'' = \{H\}$, yielding a unique rationalizable strategy profile, (H, h). As expected, this strategy profile coincides with that surviving IDSDS in example 2.2. □

Evaluating rationalizability as a solution concept

From our discussion in section 3.3, we know that, if a strategy s is rationalizable, it survives IDSDS, but the converse is not necessarily true. In particular, the set of rationalizable strategies coincides with that surviving IDSDS in two-player games, but is a subset of latter for games with three or more players. As a consequence, rationalizability satisfies the same properties as IDSDS, which holds in settings with two or more players:

1. **Existence? Yes.** Rationalizability satisfies existence, meaning that at least one strategy profile in every game must be rationalizable.

2. **Uniqueness? No.** Rationalizability does not satisfy uniqueness, since one or more strategy profiles may survive rationalizability. Although rationalizability may provide more precise equilibrium outcomes than IDSDS in games with three or more players, it does not yield a unique equilibrium prediction in all games with three or more players; entailing that neither rationalizability nor IDSDS satisfy uniqueness.

3. **Robust to small payoff perturbations? Yes.** Rationalizability is robust to small perturbations because rationalizability does not change equilibrium outcomes if we alter the payoff of one of the players by a small amount, $\varepsilon \to 0$.

4. **Socially optimal? No.** Finally, rationalizability does not necessarily yield socially optimal outcomes, as illustrated by the Prisoner's Dilemma game, where the only strategy profile surviving rationalizability is (Confess, Confess), which is not socially optimal.