Lecture G-07: Iterated Elimination of Weakly Dominated Strategies and its Relation with Subgame Perfect Equilibrium

Readings: Osborne and Rubinstein Ch 6

In this lecture we look briefly into two other related ways to think about extensive games with perfect information. We defined the procedure of iterated elimination of weakly dominated actions for a strategic game and argued that though it is less appealing than the procedure of iterated elimination of strictly dominated actions (since a weakly dominated action is a best response to *some* belief), it is a natural method for a player to use to simplify a game. In the proof of Kuhn's theorem we define the procedure of backwards induction for finite extensive games with perfect information and show that it yields the set of subgame perfect equilibria of the game.

The two procedures are related. Let Γ be a finite extensive game with perfect information in which no player is indifferent between any two terminal histories. Then Γ has a unique subgame perfect equilibrium. Let G be the strategic game induced from Γ.

In this section we will define a_i to de weakly dominated action in a strategic game G if there is an action $b_i \in A_i$ such that either:

(i) for all a_{-i} $(b_i, a_{-i}) \succsim_i (a_i, a_{-i})$ and for at least one a_{-i} we have $(b_i, a_{-i}) \succ_i (a_i, a_{-i})$.

(ii) for all a_{-i} and for all j $(b_i, a_{-i}) \sim_j (a_i, a_{-i})$ (that is b_i and a_i are equivalent from the perspective of the information buried in the game).

We now define a sequence for eliminating weakly dominated actions in the strategic form G of Γ. We will allow also deleting a strategy if there is another strategy such that whatever are the strategies of the other players it induces the same outcome. We will see that the action profiles of G that remain at the end of the procedure generate the unique subgame perfect equilibrium outcome of Γ.

Let h be a history of Γ with $P(h) = i$ and $\ell(\Gamma(h)) = 1$ and let $a_i^* \in A(h)$ be the unique action selected by the procedure of backwards induction for the history h. Backwards induction eliminates every strategy of player i that chooses an action different from a_i^* after the history h. Among these strategies, those consistent with h (i.e. that choose the component of h that follows h' whenever h' is a subhistory of h with $P(h') = i$) are weakly dominated actions in G in the above sense.

Having performed this elimination for each history h with $\ell(\Gamma(h)) = 1$, we turn to histories h with $\ell(\Gamma(h)) = 2$ and perform an analogous elimination; we continue back to the beginning of the game in this way.

Every strategy of player i that remains at the end of this procedure chooses the action selected by backwards induction after any history that is consistent with player i's subgame perfect equilibrium strategy.

In the game

1—–A—2—C—1—E—2,0
| B |D |F
3,3   1,1   0,2

The strategic game is

	C	D
AE	2,0	1,1
AF	0,2	1,1
BE	3,3	3,3
BF	3,3	3,3

In the above chain of eliminations we start with eliminating *AF* and *BF* and continue to delete *C* and finally delete *AE*. The remaining strategies are the SPE, *BE* and *D*. But one might start the elimination with eliminating *AE* following with *D*. Thus the elimination might remove all subgame perfect equilibria.

Note also that if some player is indifferent between two terminal histories then there are games where there is an order of elimination that eliminates a subgame perfect equilibrium outcome and for other games no order of elimination for which all surviving strategy profiles generate subgame perfect equilibrium outcomes.

Forward Induction

We now present two examples that show that the procedure of iterated elimination of weakly dominated strategies captures some interesting features of players' reasoning in extensive games.

Example: BoS with an outside option

Consider the extensive game with perfect information and simultaneous moves in which player 1 first decides whether to stay at home and read a book (*H*) or to try to meet player 2 (*M*). If 1 decides to read a book then the game ends; if chooses *M* the two players are engaged in the game BoS. That is, after the history (*M*) the players choose actions simultaneously. Each player prefers to meet in his favorite place (Player 1 prefers *B* and player 2 prefers *S*.

The following is the reduced strategic form of this game

	B	S
H	$2,2$	$2,2$
B	$3,1$	$0,0$
S	$0,0$	$1,3$

The only order of elimination which we can pursue is the following:

Player 1's action *S*

player 2's action *S*

Player 1's action *B*

We are left with the outcome (B,B).

This sequence of eliminations corresponds to the following natural argument referred to in the literature as "forward induction". If player 2 has to make a decision he knows that player 1 has not chosen *H*. Such a choice makes sense for player 1 only if he plans to choose *B*. Thus player 2 should choose *B* also.

Example

Two individuals are going to play BoS with monetary payoffs as in the previous example.

Before doing so player 1 can discard a dollar (take the action *D*) or refrain from doing so (take the action 0); her move is observed by player 2. Both players are risk-neutral. (Note that the two subgames that follow player 1's initial move are strategically identical.

The reduced strategic form of the game is

	BB	*BS*	*SB*	*SS*
0*B*	3, 1	3, 1	0, 0	0, 0
0*S*	0, 0	0, 0	1, 3	1, 3
DB	2, 1	−1, 0	2, 1	−1, 0
DS	−1, 0	0, 3	−1, 0	0, 3

Weakly dominated actions can be eliminated iteratively as follows.

DS is weakly dominated for player 1 by 0*B*

SS is weakly dominated for player 2 by *SB*

BS is weakly dominated for player 2 by *BB*

0*S* is strictly dominated for player 1 by *DB*

SB is weakly dominated for player 2 by *BB*

DB is strictly dominated for player 1 by 0*B*

The single strategy pair that remains is $(0B, BB)$: the fact that player 1 can throw away a dollar implies, under iterated elimination of weakly dominated actions, that the outcome is player 1's favorite.

An intuitive argument that corresponds to this sequence of eliminations is the following. player 1 must anticipate that if she chooses 0 then she will obtain an expected payoff of at least 3/4, since for every belief about the behavior of player 2 she has an action that yields her at least this expected payoff.

Thus if player 2 observes that player 1 chooses *D* then he must expect that player 1 will subsequently choose *B* (since the choice of *S* cannot possibly yield player 1 a payoff in excess of 3/4. Given this, player 2 should choose *B* if player 1 chooses *D*; player 1 knows this, so that she can expect to obtain a payoff of 2 if she chooses D. But now player 2 can rationalize the choice 0 by player 1 only by believing that player 1 will choose *B* (since *S* can yield player 1 no more than 1), so that the best action of player 2 after observing 0 is *B*. This makes 0 the best action for player 1.

This argument here is somewhat weaker than in the previous example. In BoS the action *M* is clearly irrational for player 1 if he intends to play S (and thus player 2 can conclude that he intends to play *B*). In this example, the action *D* may be rationalized by the expectation of player 1 that if he plays 0 he would get only a payoff of 1 whereas if he chooses. If player 1 does in fact expect to obtain the payoff 3 if she chooses 0, then the conclusion that player 2 should draw after observing player 1 choose *D* is unclear: it seems that player 2 may just as well conclude in this case that player 1 is irrational as conclude that player 1 will choose *B*.

From the point of view of the steady state interpretation the two examples share the same argument: the beliefs of player 2 in the equilibria in which the outcome is *H* in the first example or $(0,(S,S))$ in the second example are both unreasonable in the sense that if player 1 deviates (to *Concert* or *D*) then the only sensible conclusion for player 2 to reach is that player 1 intends to play *B* subsequently, which means that player 2 should play *B*, making the deviation profitable for player 1.

From the point of view of the deductive interpretation the two games differ, at least to the extent that the argument in the second example is more complex. In the first example player 1 has to reason about how player 2 will interpret an action (*Concert*) that she takes. In the second example player 1's reasoning about player 2's interpretation of her intended action 0 involves her belief about how player 2 would rationalize an action (*D*) that she does *not* take.

The second example raises a question about how to specify a game that captures a given situation. The arguments we have made are obviously based on the supposition that the game reflects the situation as perceived by the players. In particular, they presume that the players perceive the possibility of disposing of a dollar to be relevant to the play of BoS. Is it a plausible presumption? No reasonable person would consider the possibility of disposing a dollar to be relevant to the choice of which concert to attend. Thus we argue that a game that models the situation should simply exclude the possibility of disposal.

I would even argue that this to be so even if the game, including the move in which player 1 can burn money, is presented explicitly to the players by a referee. This is because in the first stage of the analysis of a game a player "edits" the description of the situation, eliminating "irrelevant" factors. On what principles do we base the claim that the possibility of disposing of a dollar is irrelevant? The answer is far from clear; some ideas follow. (a) The disposal does not affect the players' payoffs in BoS. (b) If the disposal is informative about the rationality of player 1, a sensible conclusion might be that a player who destroys a dollar is simply irrational. (c) The dissimilarity between the two parts of the game makes it unlikely that player 2 will try to deduce from player 1's behavior in the first stage how she will behave in the second stage.

Problem set G07

1. (**Exercise**) Describe in the extensive game tools a situation where a before a game starts one of the players can send a (free cost) message ("cheap talk" in our jargon) to player 2. Explain the sense in which this modification does not make a difference to the analysis of the game.

2. (**Exercise**) Examine a game where each player before playing the game (player 1 first and player 2 second) has the option of burning a dollar. After observing this moves they are engaged in BoS. Find the set of outcomes that survive iterated elimination of weakly dominated actions and compare it with the outcome that does so in the example above where only player 1 is allowed to burn money.