Extensive-Form Games: NE, SPE, and Backward Induction

Harry R. Lloyd, August 1, 2019

We use the extensive-form framework to model **dynamic games:** games in which players choose actions sequentially rather than simultaneously.

1 Setup

players	these can include nature / chance
history	a sequence of moves up to a given point in time
information partition[1]	a partition of the set of histories
information sets	elements of the information partition
payoffs	these are assigned to terminal histories

As always, we assume that the structure of the game is common knowledge.

1.1 Subgames

Δ is a subgame of Γ iff Δ is a continuation of Γ after a certain history such that no information set is broken up.

1.2 Actions vs. Strategies

Actions are the choices available to players (or nature) at non-terminal histories.

A strategy is a *plan of action plus a fully specified contingency plan*, listing a move for each of the player's information sets, *including those that will never be realised if the player follows the actions specified by the strategy.*

[1]A point to be clear about in the interpretation of information partitions: a player's information partition reflects the information obtained from her observations of the other players' actions *during a single play of the game*. Her experience playing the game against multiple identical opponents might give her *more information.*

1.3 Normal and Reduced Normal Form

In the normal (i.e. strategic) form, each strategy of the extensive-form game is represented by an action.
(In a simple two stage game, for player 2 'XY' might be used to abbreviate the strategy: 'play X if P1 plays a_1, and play Y is P1 plays a_2'.)

In the reduced normal form, equivalent strategies are represented by a single action. This should always be preferred to the normal form.

2 Nash Equilibrium

The Nash equilibria of an extensive-form game are the Nash equilibria of the game in normal form.

In order for players to *learn* each other's *strategies* (as the standard interpretation requires), players will either occasionally either have to deliberately deviate from their actions, or make mistakes.

Under such circumstances, an NE embodies a social norm that will be stable if each player considers changing her strategy only at the start of each game she plays, and only with regard to her *experience* of the other players' typical actions.

3 Subgame Perfect Equilibrium

A strategy profile is an SPE of a game iff it is a Nash equilibrium in every subgame.

As such, each strategy in an SPE is optimal – when the other players play as the SPE specifies – at every decision-making point where the player has full information.

In order for players to *learn* each other's *strategies* (as the standard interpretation requires), players will either occasionally either have to deliberately deviate from their actions, or make mistakes.

Under such circumstances, an SPE embodies a social norm that will be stable if each player considers changing her strategy exactly and only when she has full information, and only with regard to her *experience* of the other players' typical actions.

3.1 Existence of SPE

If (1) there are finitely many terminal histories,
and (2) every terminal history is finite, ('finite time horizon')
and (3) there is perfect information,
then: there exists at least one SPE.

3.2 Backward Induction and SPE

If (1) every terminal history is finite, ('finite time horizon')
and (2) there is perfect information,
then: the set of SPE is equal to the set of strategy profiles isolated by the procedure of backward induction.

Note that backward induction is analogous to the strategic games solution concept of *rationalizability*: it makes a – possibly broad – prediction about the behaviour of players who analyse what choices would be rational for their opponents, despite having no experience of playing the game. SPE, by contrast, is based on the idea that players *know* their opponents' actions through experience, and *reason* (in light of this knowledge) exactly and only about the optimality of *their own* actions. As such, it's to some extent surprising that these two notions coincide in a certain class of games.

3.3 Finding SPE in game with perfect information

The easiest way to do this is to use backward induction.

3.4 Finding SPE in games with imperfect information

It will often be easiest to find the NE first, and then consider which NE qualify as SPE.

It might also be possible to find the set of possible NE for particular subgames, and then apply a 'by cases' version of backward induction.

3.5 Refining SPE

We should like to require that each player's strategy be optimal – when the other players play as specified – at each of her information sets. Implementing this is less straightforward than implementing SPE, because the optimality of an action at an information set may depend on the history that has occurred, which may be unknown to the player facing the choice.

To get around this issue, we have to introduce *beliefs* into our model explicitly; for NE and SPE, we've been able to leave implicit the requirement that each player's beliefs about the other players' strategies be correct.

4 Backward Induction

Backward induction isolates play that is compatible with it being common knowledge that all *future* choices will be rational at the time when they are made.