Dynamic Games of Complete Information

Extensive-Form Games

Imperfect-Information

Mixed and Behavioral Strategies

Games with Chance Nodes

Mixed and Behavioral Strategies

Definition 68

A *mixed strategy* σ_i of player i in G_{imp} is a mixed strategy of player i in the corresponding strategic-form game $\bar{G}_{imp} = (N, (S_i)_{i \in N}, u_i)$.

Do not forget that now $s_i \in S_i$ iff s_i is a pure strategy that assigns the same action to all nodes of every information set. Hence each $s_i \in S_i$ can be seen as a function $s_i : I_i \to A$.

As before, we denote by Σ_i the set of all mixed strategies of player i and by Σ the set of all mixed strategy profiles $\Sigma_1 \times \cdots \times \Sigma_n$.

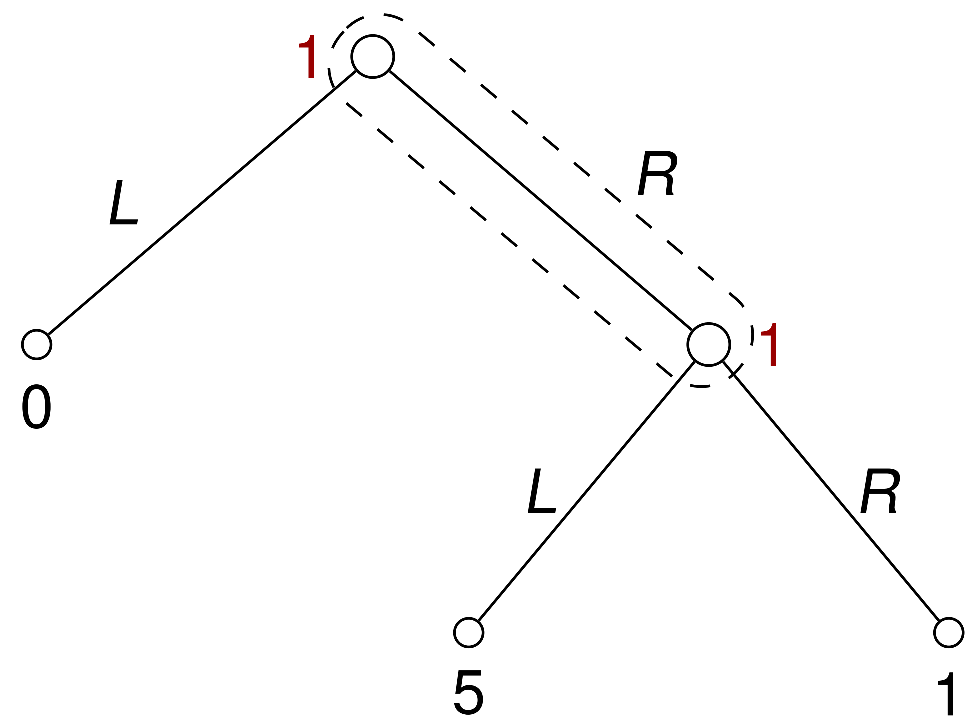
Definition 69

A *behavioral strategy* of player i in G_{imp} is a behavioral strategy β_i in G_{perf} such that for all $j = 1, \ldots, k_i$ and all $h, h' \in I_{i,j} : \beta_i(h) = \beta_i(h')$.

Each β_i can be seen as a function $\beta_i : I_i \to \Delta(A)$ such that for all $I_{i,j} \in I_i$ we have $supp(\beta_i(I_{i,j})) \subseteq \chi(I_{i,j})$.

Are they equivalent as in the perfect-information case?

Example: Absent Minded Driver



Only one player: A driver who has to take a turn at a particular junction. There are two identical junctions, the first one leads to a wrong neighborhood where the driver gets completely lost (payoff 0), the second one leads home (payoff 5). If the driver misses both, there is a longer way home (payoff 1). The problem is that after missing the first turn, the driver forgets that he missed the turn.

Behavioral strategy: $\beta_1(I_{1,1})(L) = \frac{1}{2}$ has the expected payoff $\frac{3}{2}$.

No mixed strategy gives a larger payoff than 1 since no pure strategy ever reaches the terminal node with payoff 5.

Kuhn's Theorem

Definition 70
Player i has *perfect recall* in G_{imp} if the following holds:

- Every information set of player i intersects every path from the root h_0 to a terminal node at most once.
- Every two paths from the root that end in the same information set of player i
  - pass through the same information sets of player i,
  - and in the same order,
  - and in every such information set the two paths choose the same action.

In other words, along all paths ending in the same information set, player i sees the same sequence of information sets and makes the same decisions in his nodes (i.e. at the end knows exactly the sequence of visited information sets and all his own choices along the way).

Kuhn's Theorem

The notion of induced strategies can be straightforwardly generalized to imperfect information games:

Behavioral to mixed: We say that a mixed strategy σ_i is *induced by* a behavioral strategy β_i if

$$\sigma_i(s_i) = \prod_{I_{i,j} \in I_i} \beta_i(I_{i,j})(s_i(I_{i,j})) \text{ for all } s_i \in S_i$$

As before, for the opposite direction some notation is needed. Recall that given $h \in \mathcal{H}$, we denote by $w[h]$ the unique path from h_0 to h.

Given $h \in H_i$, we denote by S_i^h the set of all pure strategies $s_i \in S_i$ such that for every $h' \in H_i$ visited by $w[h]$ we have that $s_i(h')$ is the action chosen in h' on $w[h]$.

Given $h \in H_i$ and $a \in \chi(h)$, we denote by $S_i^{h,a}$ the set of all pure strategies $s_i \in S_i^h$ such that $s_i(h) = a$.

Kuhn's Theorem

Mixed to behavioral: We say that a behavioral strategy β_i is *induced by* a mixed strategy σ_i if the following holds:

Let $I_{i,j}$ be an information set of player i and let $h \in I_{i,j}$ be (an arbitrary) node of $I_{i,j}$. We have that

- either $\sum_{s_i \in S_i^h} \sigma_i(s_i) = 0$
- or for each action $a \in \chi(h)$ $(= \chi(I_{i,j}))$:

$$\beta_i(I_{i,j})(a) = \frac{\sum_{s_i \in S_i^{h,a}} \sigma_i(s_i)}{\sum_{s_i \in S_i^h} \sigma_i(s_i)}$$

(Here the perfect recall implies that the definition of $\beta_i(I_{i,j})$ does not depend on the choice of h.)

Theorem 71 (Kuhn, 1953)

Let α be a mixed/behavioral strategy profile and let α' be any mixed/behavioral profile obtained from α by substituting some of the strategies in α with strategies they induce. Then $u_i(\alpha) = u_i(\alpha')$.

The concepts of Nash equilibria and SPE in behavioral strategies are the same as in the perfect information case.

Complexity of Zero-Sum Games

Recall that a behavioral strategy β_i of player i is maxmin if

$$\beta_i \in \operatorname*{argmax}_{\beta'_i \in \mathcal{B}_i} \min_{\beta_{-i} \in \mathcal{B}_{-i}} u_i(\beta'_i, \beta_{-i})$$

Similarly for pure and mixed strategies.

Theorem 72 (Koller and Megiddo, 1990)

Consider finite two-player zero-sum imperfect information games.

- *For such games with perfect recall, the problem of computing a maxmin behavioral strategy is in PTIME.*
- *For games with possibly imperfect recall, the problem of computing a (pure, behavioral, or mixed) strategy that guarantees a given payoff is NP-hard.*

How to compute Nash equilibria in polynomial time?

Existence of a poly. time algorithm for computing behavioral NE does not immediately follow from Thm 72 and von Neumann's Thm 48. Indeed, Thm 48 has been proved only for mixed strategies. However, using Kuhn's thm, von Neumann's thm can be easily extended to behavioral strategies.

Complexity of Zero-Sum Games

Proposition 5

Let (β_1, β_2) be a behavioral strategy profile. Then (β_1, β_2) is a NE iff both β_1 and β_2 are maxmin.

Proof. Let (β_1, β_2) be a profile of behavioral strategies. Apply Kuhn's theorem and obtain induced mixed strategies (σ_1, σ_2).

Since we used only the Kuhn's theorem to obtain (σ_1, σ_2) from (β_1, β_2), for both $i \in \{1, 2\}$ holds: $u_i(\beta_1, \beta_2) = u_i(\sigma_1, \sigma_2)$ and

- for every behavioral strategy β'_{-i} and an induced mixed strategy σ'_{-i}, we have that $u_i(\beta_i, \beta'_{-i}) = u_i(\sigma_i, \sigma'_{-i})$,
- for every mixed strategy σ'_{-i} and an induced behavioral strategy β'_{-i}, we have that $u_i(\sigma_i, \sigma'_{-i}) = u_i(\beta_i, \beta'_{-i})$.

Now (β_1, β_2) is a Nash equilibrium iff (σ_1, σ_2) is a Nash equilibrium iff σ_1 and σ_2 are maxmin iff β_1 and β_2 are maxmin.

Corollary 73

The complexity of computing Nash equilibria in behavioral strategies in two-player zero-sum imperfect information games with perfect recall is in PTIME.

Complexity of Non-Zero-Sum Games

Computing NE (or SPE) in non-zero-sum imperfect-information extensive-form games is at least as hard as for strategic-form games.

Backward induction helps in decomposing the game into "subgames" rooted in nodes of H_{single} but large games may still remain to be solved using other methods.

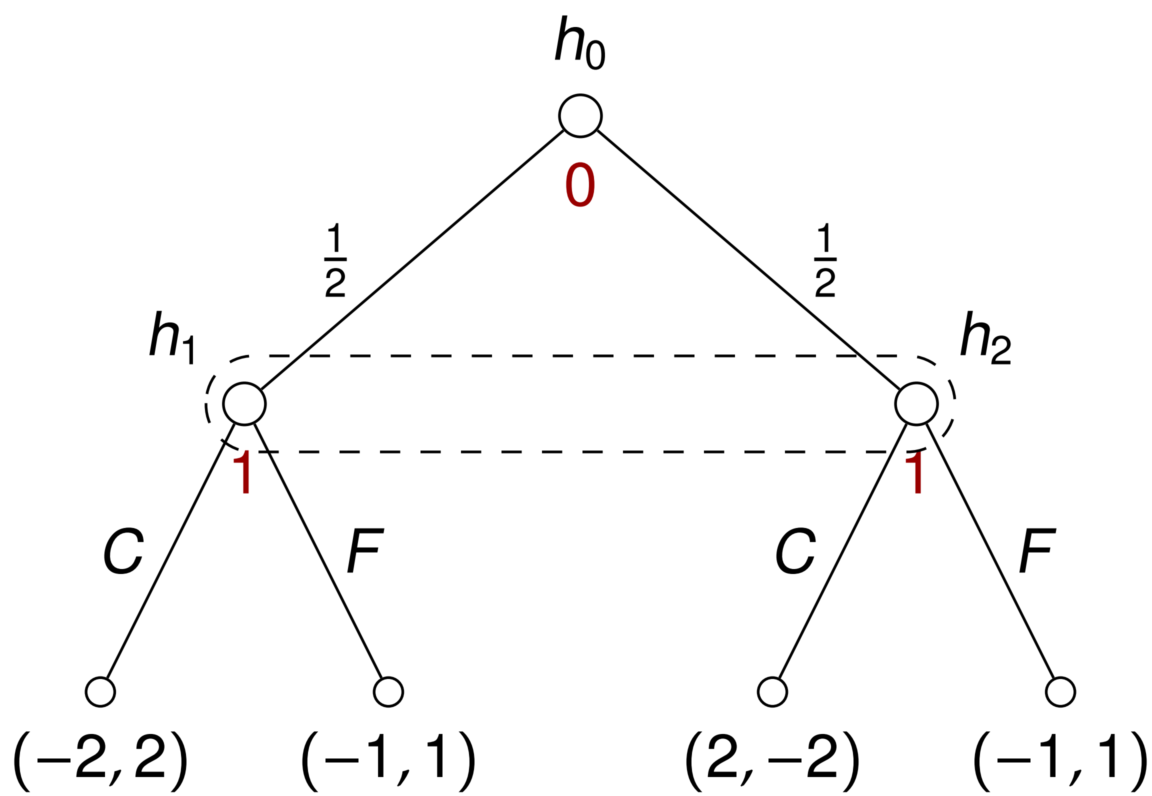
Naively, any solution concept developed for strategic-form games can be applied to imperfect-information extensive-form games (with perfect recall) via the corresponding strategic-form game $\bar{G}_{imp}$.

However, such solution is not efficient (the corresponding game is exponentially large and often degenerate).

More efficient methods exist for two-player games of perfect recall, e.g., using *sequence form representation* of the game, where nodes of G_{imp} are represented by sequences of actions leading to the nodes, which leads to a linear complementarity problem of *polynomial size*, which in turn can be solved using a modified Lemke-Howson.

For a detailed treatment of complexity see "The complexity of computing a (perfect) equilibrium for an n-player extensive form game of perfect recall" by Kousha Etessami.

Imperfect-information and Chance Nodes



A very simple card game:

- Player 1 chooses randomly a card from a large deck of cards, containing only an equal number of Kings and Aces.
- Then Player 1 may either call (C) or fold (F), no look at the card.
- If he folds, then pays $1 to player 2, otherwise
  - call + King means that player 1 pays $2 to player 2
  - call + Ace means that player 2 pays $2 to player 1

Imperfect-information and Chance Nodes

An *imperfect-information extensive-form game with chance nodes* is a tuple $G_{imp} = (G_{perf}, I, \beta_0)$ where

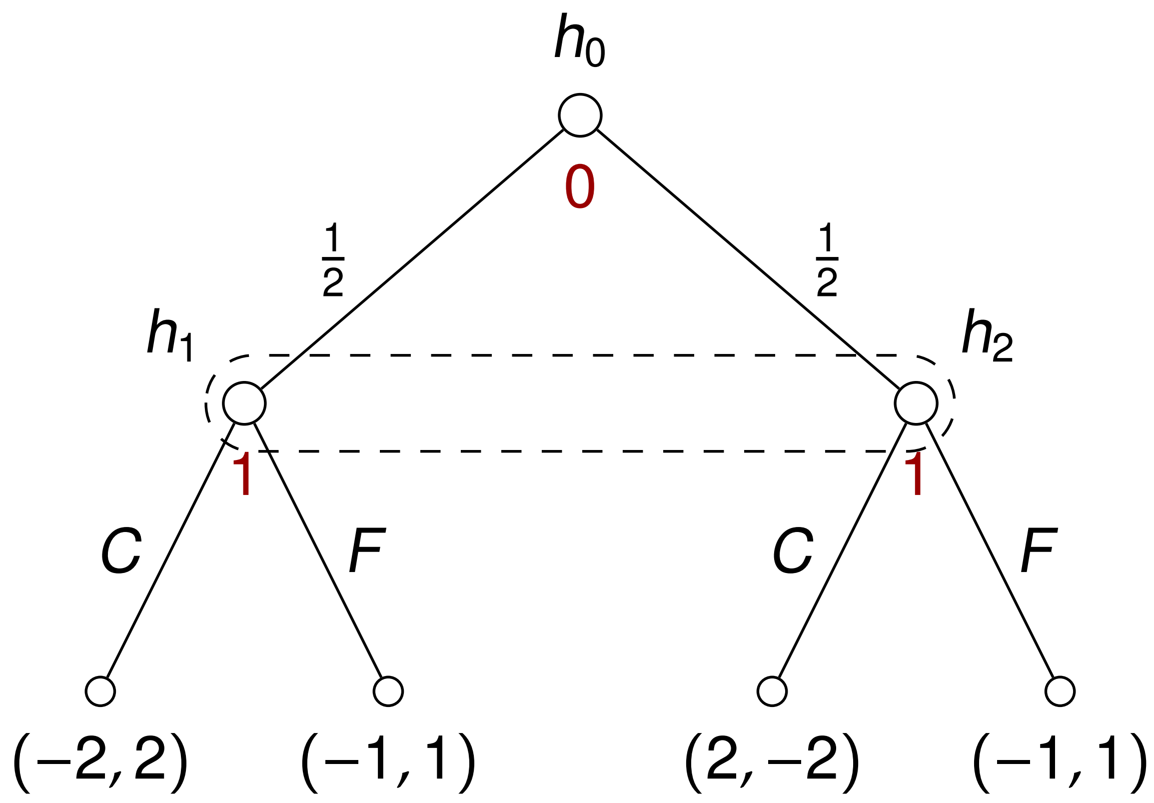
- The set of players N is equal to $\{0, 1, \dots, n\}$ (i.e., there is a new player 0 called *chance*, or *nature*),
- We assume that for every $h \in H_0$ the set of enabled actions $\chi(h)$ is the set of all children nodes of h,
- Each information set of player 0 is a singleton (i.e., the nature has a complete information),
- β_0 is a fixed *behavioral* strategy for player 0. Player 0 *always* plays according to β_0.

  Note that due to the above assumption, $\beta_0(h)$ is a distribution on all children of h

As player 0 plays the same strategy always, we exclude this strategy from strategy profiles
(i.e. pure strategy profiles remain to be elements of $S_1 \times \cdots \times S_n$)

A game with chance nodes is a *perfect information game* if all information sets of I are singletons.

Example



Here $\beta_0(h_0)(h_1) = \frac{1}{2}$ and $\beta_0(h_0)(h_2) = \frac{1}{2}$

Player 1 has just one information set $I_{1,1} = \{h_1, h_2\}$.

Consider a mixed strategy σ_1 of player 1 defined by $\sigma_1(I_{1,1})(C) = \frac{1}{4}$ and $\sigma_1(I_{1,1})(F) = \frac{3}{4}$.

Then $u_1(\sigma_1) = \frac{1}{2}\frac{1}{4}(-2) + \frac{1}{2}\frac{3}{4}(-1) + \frac{1}{2}\frac{1}{4}2 + \frac{1}{2}\frac{3}{4}(-1) = -\frac{3}{4}$

Results

All results for games without chance nodes presented so far remain valid for games with chance nodes.

In particular, Theorem 57 and Theorem 64 remain valid for games of perfect information with chance nodes. Concretely:

Theorem 74

Consider games of perfect information with chance nodes.

- *There exists a pure strategy profile which is a SPE with respect to pure strategies.*
- *There exists a pure strategy profile which is a SPE with respect to behavioral strategies.*

Backward induction can be straightforwardly modified to deal with chance nodes (see next slide).

Backward induction with perfect info. & chance

Backward Induction: We inductively "attach" to every node h a SPE s^h in G^h, and expected payoffs $u(h) = (u_1(h), \ldots, u_n(h))$.

- **Initially:** Attach to each terminal node $z \in Z$ the empty profile $s^z = (\emptyset, \ldots, \emptyset)$ and the payoff vector $u(z) = (u_1(z), \ldots, u_n(z))$.
- **While**(there is an unattached node h with all children attached):
  1. Let K be the set of all children of h
  2. If $\chi(h) \neq 0$ then let $h_{max} \in \text{argmax}_{h' \in K}\, u_{\rho(h)}(h')$ and
     - attach to h a SPE s^h where $s^h_{\rho(h)}(h) = h_{\max}$ and for $i \in N \setminus \{0\}$ and all $h' \in H_i$ define $s^h_i(h') = s^{\bar{h}}_i(h')$ where $h' \in H^{\bar{h}} \cap H_i$

       (i.e. in subgames rooted in $\bar{h} \in K$, s^h behaves as $s^{\bar{h}}$.)
     - attach to h expected payoffs $u_i(h) = u_i(h_{\max})$ for $i \in N \setminus \{0\}$
  3. If $\chi(h) = 0$, then
     - attach to h a SPE s^h where for all $i \in N \setminus \{0\}$ and all $h' \in H_i$ define $s^h_i(h') = s^{\bar{h}}_i(h')$ where $h' \in H^{\bar{h}} \cap H_i$

       (i.e. in subgames rooted in $\bar{h} \in K$, s^h behaves as $s^{\bar{h}}$.)
     - attach to h the expected payoffs

$$u_i(h) = \sum_{\bar{h} \in K} \big(\beta_0(h)(\bar{h})\big)\, u_i(\bar{h})$$

(i.e., the weighted average payoff in all children nodes)

Backward Induction for Imperfect Info & Chance

The high-level description of backward induction for imperfect-information games given earlier remains valid also for imperfect-information games with chance nodes.

We only have to notice that in the games newly created in step 4., player 0 participates with the strategy β_0.

Dynamic Games of Complete Information

Repeated Games

Finitely Repeated Games

Example

	C	S
C	$-5, -5$	$0, -20$
S	$-20, 0$	$-1, -1$

Imagine that the criminals are being arrested repeatedly.

Can they somewhat reflect upon their experience in order to play "better"?

In what follows we consider strategic-form games played repeatedly

- for finitely many rounds, the final payoff of each player will be the average of payoffs from all rounds
- infinitely many rounds, here we consider a discounted sum of payoffs and the long-run average payoff

We analyze Nash equilibria and sub-game perfect equilibria.

We stick to pure strategies only!

Finitely Repeated Games

Let $G = (\{1,2\},(S_1,S_2),(u_1,u_2))$ be a finite strategic-form game of two players.

A *T-stage game $G_{T\text{-}rep}$ based on G* proceeds in T *stages* so that in a stage $t \geq 1$, players choose a strategy profile $s^t = (s_1^t, s_2^t)$.

After T stages, both players collect the average payoff $\sum_{t=1}^{T} u_i(s^t) \,/\, T$.

A *history of length $0 \leq t \leq T$* is a sequence $h = s^1 \cdots s^t \in S^t$ of t strategy profiles. Denote by $H(t)$ the set of all histories of length t.

A *pure strategy* for player i in a T-stage game $G_{T\text{-}rep}$ is a function

$$\tau_i : \bigcup_{t=0}^{T-1} H(t) \to S_i$$

which for every possible history chooses a next step for player i.

Every strategy profile $\tau = (\tau_1, \tau_2)$ in $G_{T\text{-}rep}$ induces a sequence of pure strategy profiles $w_\tau = s^1 \cdots s^T$ in G so that $s_i^t = \tau_i(s^1 \cdots s^{t-1})$.

Given a pure strategy profile τ in $G_{T\text{-}rep}$ such that $w_\tau = s^1 \cdots s^T$, define the payoffs $u_i(\tau) = \sum_{t=1}^{T} u_i(s^t) \,/\, T$.

Example

	C	S
C	$-5, -5$	$0, -20$
S	$-20, 0$	$-1, -1$

Consider a 3-stage game.

Examples of histories: ϵ, (C, S), $(C, S)(S, S)$, $(C, S)(S, S)(C, C)$

Here the last one is terminal, obtained using τ_1, τ_2 s.t.:

$\tau_1(\epsilon) = C$, $\tau_1((C, S)) = S$, $\tau_1((C, S)(S, S)) = C$
$\tau_2(\epsilon) = S$, $\tau_2((C, S)) = S$, $\tau_2((C, S)(S, S)) = C$
Thus $w_{(\tau_1, \tau_2)} = (C, S)(S, S)(C, C)$

$u_1(\tau_1, \tau_2) = (0 + (-1) + (-5))/3 = -2$
$u_2(\tau_1, \tau_2) = (-20 + (-1) + (-5))/3 = -26/3$

Finitely Repeated Games in Extensive-Form

Every T-stage game $G_{T\text{-}rep}$ can be defined as an imperfect information extensive-form game.

Define an imperfect-information extensive-form game $G^{rep}_{imp} = (G^{rep}_{perf}, I)$ such that $G^{rep}_{perf} = (\{1, 2\}, A, H, Z, \chi, \rho, \pi, h_0, u)$ where

- $A = S_1 \cup S_2$
- $H = (S_1 \times S_2)^{\leq T} \cup (S_1 \times S_2)^{<T} \cdot S_1$

  Intuitively, elements of $(S_1 \times S_2)^{\leq k}$ are possible histories; $(S_1 \times S_2)^{<k} \cdot S_1$ is used to simulate a simultaneous play of G by letting player 1 choose first and player 2 second.
- $Z = (S_1 \times S_2)^T$
- $\chi(\epsilon) = S_1$ and $\chi(h \cdot s_1) = S_2$ for $s_1 \in S_1$, and $\chi(h \cdot (s_1, s_2)) = S_1$ for $(s_1, s_2) \in S$
- $\rho(\epsilon) = 1$ and $\rho(h \cdot s_1) = 2$ and $\rho(h \cdot (s_1, s_2)) = 1$
- $\pi(\epsilon, s_1) = s_1$ and $\pi(h \cdot s_1, s_2) = h \cdot (s_1, s_2)$ and $\pi(h \cdot (s_1, s_2), s_1') = h \cdot (s_1, s_2) \cdot s_1'$
- $h_0 = \epsilon$ and $u_i((s_1^1, s_2^1)(s_1^2, s_2^2) \cdots (s_1^T, s_2^T)) = \sum_{t=1}^{T} u_i(s_1^t, s_2^t) \,/\, T$

Finitely Repeated Games in Extensive-Form

The set of information sets is defined as follows: Let $h \in H_1$ be a node of player 1, then

- there is exactly one information set of player 1 containing h as the only element,
- there is exactly one information set of player 2 containing all nodes of the form $h \cdot s_1$ where $s_1 \in S_1$.

Intuitively, in every round, player 1 has a complete information about results of past plays,

player 1 chooses a pure strategy $s_1 \in S_1$,

player 2 is *not* informed about s_1 but still has a complete information about results of all previous rounds,

player 2 chooses a pure strategy $s_2 \in S_2$ and both players are informed about the result.

Finitely Repeated Games – Equilibria

Definition 75
A strategy profile $\tau = (\tau_1, \tau_2)$ in a T-stage game $G_{T\text{-}rep}$ is a Nash equilibrium if for every $i \in \{1, 2\}$ and every τ_i' we have

$$u_i(\tau_1, \tau_2) \geq u_i(\tau_i', \tau_{-i})$$

To define SPE we use the following notation. Given a history $h = s^1 \cdots s^t$ and a strategy τ_i of player i, we define a strategy τ_i^h in $(T - t)$-stage game based on G by

$$\tau_i^h(\bar{s}^1 \cdots \bar{s}^{\bar{t}}) = \tau_i(s^1 \cdots s^t \bar{s}^1 \cdots \bar{s}^{\bar{t}}) \quad \text{for every sequence } \bar{s}^1 \cdots \bar{s}^{\bar{t}}$$

(i.e. τ_i^h behaves as τ_i after h)

Definition 76
A strategy profile $\tau = (\tau_1, \tau_2)$ in a T-stage game $G_{T\text{-}rep}$ is a subgame-perfect Nash equilibrium (SPE) if for every history h the profile (τ_1^h, τ_2^h) is a Nash equilibrium in the $(T - |h|)$-stage game based on G.

SPE with Single NE in G

	C	S
C	$-5, -5$	$0, -20$
S	$-20, 0$	$-1, -1$

Consider a T-stage game based on Prisoner's dilemma.

For every T, find a SPE.

... there is one, play (C, C) all the time. Is it all?

Theorem 77
Let G be an arbitrary finite strategic-form game. If G has a unique Nash equilibrium, then playing this equilibrium all the time is the unique SPE in the T-stage game based on G.

Proof.
By backward induction, players have to play the NE in the last stage. As the behavior in the last stage does not depend on the behavior in the $(T-1)$-th stage, they have to play the NE also in the $(T-1)$-th stage. Then the same holds in the $(T-2)$-th stage, etc. □

Further Discussion of Prisoner's Dilemma

	C	S
C	$-5, -5$	$0, -20$
S	$-20, 0$	$-1, -1$

Are there other NE (that are not SPE) in the repeated Prisoner's dilemma?

To simplify our discussion, we use the following notation: $X{-}YZ$, where $X, Y, Z \in \{C, S\}$ denotes the following strategy:

- In the first phase, play X
- In the second phase, play Y if the opponent plays C in the first phase, otherwise play Z

There are 4 NE: They are the four profiles that lead to $(C, C)(C, C)$, i.e., each player plays either $C{-}CC$, or $C{-}CS$.

Further Discussion of Prisoner's Dilemma

	C	S
C	$-5, -5$	$0, -20$
S	$-20, 0$	$-1, -1$

The strategy C strictly dominates S in the Prisoner's dilemma.

Is there a strictly dominant strategy in the 2-stage game based on the Prisoner's dilemma?

If player 2 plays $S{-}CS$, then the best responses of player 1 are $S{-}CC$ and $S{-}SC$.
(The strategy $S{-}CS$ is usually called "tit-for-tat".)

If player 2 plays $S{-}SC$, then the best responses are $C{-}SC$ and $C{-}CC$.

So there is no strictly dominant strategy for player 1.
(Which would be among the best responses for all strategies of player 2.)

SPE with Multiple NE in G

Let $s = (s_1, s_2)$ be a Nash equilibrium in G.

Define a strategy profile $\tau = (\tau_1, \tau_2)$ in $G_{T\text{-}rep}$ where

- τ_1 chooses s_1 in every stage
- τ_2 chooses s_2 in every stage

Proposition 6

τ is a SPE in $G_{T\text{-}rep}$ for every $T \geq 1$.

Proof.

Apparently, changing τ_i in some stage(s) may only result in the same or worse payoff for player i, since the other player always plays s_2 independent of the choices of player 1. □

The proposition may be generalized by allowing players to play different equilibria in particular stages

I.e., consider a sequence of NE $s^1, s^2, \ldots, s^T$ in G and assume that in stage ℓ player i plays s_i^ℓ

Does this cover all possible SPE in finitely repeated games?

SPE with Multiple NE in G

	m	f	r
M	$4,4$	$-1,5$	$0,0$
F	$5,-1$	$1,1$	$0,0$
R	$0,0$	$0,0$	$3,3$

NE in the above game G : (F, f) and (R, r)

Consider 2-stage game $G_{2\text{-}rep}$ and strategies τ_1, τ_2 where

- τ_1 : Chooses M in stage 1. In stage 2 plays R if (M, m) was played in the first stage, and plays F otherwise.
- τ_2 : Chooses m in stage 1. In stage 2 plays r if (M, m) was played in the first stage, and plays f otherwise.

Is this SPE?

Note that here the players **do not** play a NE in the first step.

The idea is that both players agree to play a Pareto optimal profile. If both comply, then a favorable NE is played in the second stage. If one of them betrays then a "punishing" NE is played.