Dresden University of Technology

Computational Logic Group

Hannes Strass

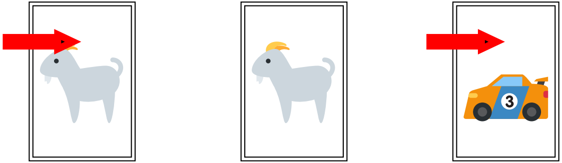
Faculty of Computer Science, Institute of Artificial Intelligence, Computational Logic Group

Games with Missing Information: Solving

Lecture 8, 8th Jun 2026 // Algorithmic Game Theory, SS 2026

Previously . . .

- In **complete information** games, players know the rules, possible outcomes and each other's preferences over outcomes.
- In **perfect information** games, all players know all previous moves.
- An **extensive-form** game may have incomplete or imperfect information.
- Uncertainty of players (due to missing information) can be modelled by **information sets** and **chance nodes** (moves by `Nature`).
- A **behaviour strategy** assigns move probabilities to information sets.
- In a game with **perfect recall**, players remember their own previous moves and the information sets in which they played them.
- **Kuhn's Theorem:** Considering only behaviour strategies "is enough" for playing games with perfect recall.



Overview

Example: Simplified Poker

Belief Systems

(Weak) Sequential Equilibria

Solving Simplified Poker

Example: Simplified Poker

Simplified Poker: Game Description

Binmore's Simplified Poker

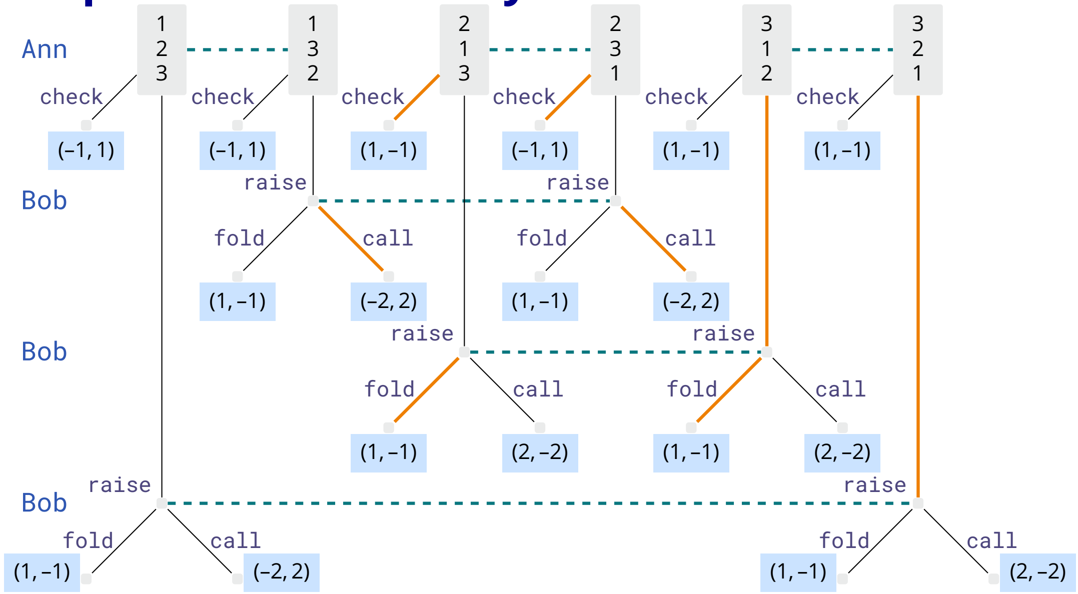
- Two players, Ann and Bob, each put \$1 into a jackpot.
- They then draw one card from a deck of three cards: $\{1, 2, 3\}$.
- Ann can either `check` (pass on), or `raise` (put another \$1 into the jackpot).
- Next, Bob responds:
  - If Ann has `check`ed, then Bob must `call`, that is, a showdown happens:
    Both players show their cards and the player with the higher (number) card receives the jackpot.
  - If Ann has `raise`d, then Bob can decide between `fold` (withdraw from the game and let Ann get the jackpot) or `call` (put another \$1 into the jackpot and then have a showdown).

Simplified Poker: Formal Model

Simplified Poker can be modelled as an extensive-form game as follows:

- $P = \{\texttt{Ann}, \texttt{Bob}, \texttt{Nature}\}$
- $\mathbf{M} = (M_{\texttt{Ann}}, M_{\texttt{Bob}}, M_{\texttt{Nature}})$ with
  - $M_{\texttt{Ann}} = \{\texttt{check}, \texttt{raise}\}$,
  - $M_{\texttt{Bob}} = \{\texttt{fold}, \texttt{call}\}$,
  - $M_{\texttt{Nature}} = \{\texttt{deal123}, \texttt{deal132}, \texttt{deal213}, \texttt{deal231}, \texttt{deal312}, \texttt{deal321}\}$.
- $\mathcal{I} = \{\mathcal{I}_{A1}, \mathcal{I}_{A2}, \mathcal{I}_{A3}, \mathcal{I}_{B1}, \mathcal{I}_{B2}, \mathcal{I}_{B3}\}$ with
  - $\mathcal{I}_{A1} = \{[\texttt{deal123}], [\texttt{deal132}]\}$,
    $\mathcal{I}_{A2} = \{[\texttt{deal213}], [\texttt{deal231}]\}$,
    $\mathcal{I}_{A3} = \{[\texttt{deal312}], [\texttt{deal321}]\}$ with $p(\mathcal{I}_{A1}) = p(\mathcal{I}_{A2}) = p(\mathcal{I}_{A3}) = \texttt{Ann}$,
  - $\mathcal{I}_{B1} = \{[\texttt{deal213}, \texttt{raise}], [\texttt{deal312}, \texttt{raise}]\}$,
    $\mathcal{I}_{B2} = \{[\texttt{deal123}, \texttt{raise}], [\texttt{deal321}, \texttt{raise}]\}$,
    $\mathcal{I}_{B3} = \{[\texttt{deal132}, \texttt{raise}], [\texttt{deal231}, \texttt{raise}]\}$ with $p(\mathcal{I}_{B1}) = p(\mathcal{I}_{B2}) = p(\mathcal{I}_{B3}) = \texttt{Bob}$.
- $\mathbf{u} = (u_{\texttt{Ann}}, u_{\texttt{Bob}})$ with the functions as shown next in the game tree.

Simplified Poker: Analysis



Simplified Poker: Open Questions

What happens in the two remaining cases?

1. Should `Ann` `raise` (i.e. bluff) if she has a 1?
2. Should `Bob` `call` (the bluff) if he has a 2?

Belief Systems

Recall: Behaviour Strategies

Example (Simplified Poker)

Consider information set $\mathcal{I}_{\mathtt{A1}} = \{[\mathtt{deal123}], [\mathtt{deal132}]\}$ where Ann has a 1. With $\sigma_{\mathtt{Ann}}(\mathcal{I}_{\mathtt{A1}}) = \left\{\mathtt{check} \mapsto \frac{1}{2}, \mathtt{raise} \mapsto \frac{1}{2}\right\}$, she bases her decision to bluff (with her 1) on a (balanced) coin flip.

A behaviour strategy profile $\boldsymbol{\sigma}$ induces an expected utility for each player $i \in P$:

$$U_i(\boldsymbol{\sigma}) := \sum_{z \in Z} P(z \,|\, \boldsymbol{\sigma}) \cdot u_i(z)$$

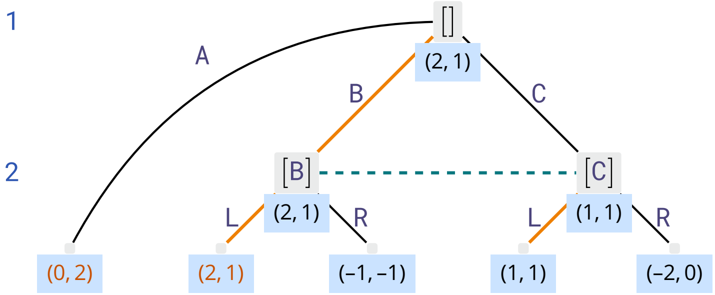
where $P(h \,|\, \boldsymbol{\sigma})$ is the probability that history h is reached whenever play happens according to profile $\boldsymbol{\sigma}$: inductively, define $P([] \,|\, \boldsymbol{\sigma}) := 1$ and

$$P([h; m] \,|\, \boldsymbol{\sigma}) := \sigma_{p(\mathcal{I}_h)}(m \,|\, \mathcal{I}_h) \cdot P(h \,|\, \boldsymbol{\sigma})$$

- where $\mathcal{I}_h \in \mathcal{I}$ is the unique information set with $h \in \mathcal{I}_h$,
- and $\sigma_{\mathtt{Nature}}$ is obtained from the probability distributions specified by G.

Towards Solution Concepts: Example

Consider the following extensive-form game G_4 and its normal form:



(1, 2)	L	R
A	(0, 2)	(0, 2)
B	(2, 1)	(−1, −1)
C	(1, 1)	(−2, 0)

- The normal form game has two pure Nash equilibria: (A, R) and (B, L).
- Arguably, only (B, L) respects sequentiality:
  - If play reaches {[B], [C]}, then 2 will choose L.
  - Knowing this, 1 will choose B.

⇝ Adapt subgame perfect equilibria to information sets?

Subgames of Extensive-Form Games

Definition

Let G be an extensive-form game. A **subgame** G' of G consists of:

- A non-terminal history $h' \in H$ of G, the **root** of G',
- all histories $H' \subseteq H$ of G that start with h' (including $Z' = H' \cap Z$), and
- all other aspects of G restricted to H' (players, moves, information sets, turn function p, probability distributions for `Nature`, and utilities),

where for all $\mathcal{I}_j \in \mathcal{I}$, either $\mathcal{I}_j \cap H' = \mathcal{I}_j$ or $\mathcal{I}_j \cap H' = \emptyset$.

Observation

If G' is a subgame of G, then its root h' is in information set $\{h'\}$.

Example

G_4 only has the trivial subgame, itself.

Towards Solution Concepts: Stocktaking

- Viewing an extensive-form game as a normal-form game, we could obtain (mixed) Nash equilibria.
- That did not fully work even for perfect-information sequential games:
- There, we used a stronger solution concept: subgame perfect equilibria, where strategies must play best responses in all subgames.
- With information sets, not every decision point corresponds to a subgame.
- Information sets off the equilibrium path might be relevant.

Example (G_4)

- G_4 has only itself as subgame, so equilibrium (A, R) is "subgame perfect".
- In (A, R), information set $\mathcal{I}_2 = \{[\mathtt{B}], [\mathtt{C}]\}$ is reached with probability zero.
- To define playing best responses "everywhere": What is 2's expected payoff from information set $\mathcal{I}_2$ when play happens as in (A, R)?

$\rightsquigarrow$ We will additionally model players' beliefs about histories . . .

Belief Systems

Definition

Let G be an extensive-form game with n players and information sets $\mathcal{I}$.

A **belief system** for G is a tuple $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$ of functions β_i that assign

- to each $\mathcal{I}_j \in \mathcal{I}$ with $p(\mathcal{I}_j) = i \neq \texttt{Nature}$
- a probability distribution $\beta_i(\mathcal{I}_j)$ on histories $h \in \mathcal{I}_j$.

- We denote $\beta_i(\mathcal{I}_j)(h)$ by $\beta_i(h \mid \mathcal{I}_j)$;
- the value $\beta_i(h \mid \mathcal{I}_j)$ reflects player i's (where $i = p(\mathcal{I}_j)$) belief about the likelihood that h is the true history, given that i knows to be in $\mathcal{I}_j$.

Example (Simplified Poker)

- In belief system β_{Ann} with $\beta_{\text{Ann}}(\mathcal{I}_{\text{A1}}) = \left\{[\texttt{deal123}] \mapsto \frac{1}{2}, [\texttt{deal132}] \mapsto \frac{1}{2}\right\}$, Ann considers "Bob has a 2" and "Bob has a 3" to be equally likely.
- If $\beta_{\text{Bob}}([\texttt{deal123}, \texttt{raise}] \mid \mathcal{I}_{\text{B2}}) = 0$, then Bob is sure that Ann does not bluff.

Assessments

Definition

Let G be an extensive-form game with non-Nature players $1, \ldots, n$.

An **assessment** of G is a pair $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ consisting of a profile $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ of behaviour strategies and a belief system $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$.

Example (Simplified Poker)

Consider the assessment $(\boldsymbol{\sigma}', \boldsymbol{\beta}')$ with

- $\sigma'_{\mathsf{Ann}}(\mathcal{I}_{\mathsf{A1}}) = \left\{\mathtt{check} \mapsto \frac{1}{2}, \mathtt{raise} \mapsto \frac{1}{2}\right\}$, and playing optimally elsewhere,
- $\sigma'_{\mathsf{Bob}}(\mathcal{I}_{\mathsf{B2}}) = \left\{\mathtt{fold} \mapsto \frac{1}{2}, \mathtt{call} \mapsto \frac{1}{2}\right\}$, and playing optimally elsewhere;
- $\beta'_{\mathsf{Ann}}(\mathcal{I}_{\mathsf{A1}})$, $\beta'_{\mathsf{Ann}}(\mathcal{I}_{\mathsf{A2}})$, and $\beta'_{\mathsf{Ann}}(\mathcal{I}_{\mathsf{A3}})$ all uniform distributions,
- where in $\mathcal{I}_{\mathsf{B3}}$ and $\mathcal{I}_{\mathsf{B1}}$ Bob is sure that Ann does not raise with a 2, and
- $\beta'_{\mathsf{Bob}}(\mathcal{I}_{\mathsf{B2}}) = \left\{[\mathtt{deal123}, \mathtt{raise}] \mapsto \frac{1}{4}, [\mathtt{deal321}, \mathtt{raise}] \mapsto \frac{3}{4}\right\}$.

Expected Utility for Assessments

Definition

Let G be an extensive-form game and $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ be an assessment of G.

The **expected utility** for player i at information set $\mathcal{I}_j$ according to $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ is

$$U_i(\mathcal{I}_j, \boldsymbol{\sigma}, \boldsymbol{\beta}) := \sum_{h \in \mathcal{I}_j} \left(\beta_i(h \mid \mathcal{I}_j) \cdot \sum_{z \in Z} \left(P(z \mid h, \boldsymbol{\sigma}) \cdot u_i(z) \right) \right)$$

where $P(h' \mid h, \boldsymbol{\sigma})$ is the probability that history h' is reached when playing according to $\boldsymbol{\sigma}$ from history h on:

$$\begin{aligned} P(h \mid h, \boldsymbol{\sigma}) &:= 1 && \text{for all } h \in H \\ P([] \mid h, \boldsymbol{\sigma}) &:= 0 && \text{for all } h \neq [] \\ P([h'; m] \mid h, \boldsymbol{\sigma}) &:= \sigma_{p(\mathcal{I}_{h'})}(m \mid \mathcal{I}_{h'}) \cdot P(h' \mid h, \boldsymbol{\sigma}) \end{aligned}$$

Example: $U_{\text{Bob}}(\mathcal{I}_{\text{B2}}, \boldsymbol{\sigma}', \boldsymbol{\beta}') = \frac{1}{4} \cdot \left(\frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot 2 \right) + \frac{3}{4} \cdot \left(\frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot (-2) \right) = -1.$

(Weak) Sequential Equilibria

Best Responses and Sequential Rationality

Definition

Let $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ be an assessment for an extensive-form game G with players P.

1. Player i's strategy σ_i is a **best response** to $\boldsymbol{\sigma}_{-i}$ w.r.t. $\boldsymbol{\beta}$ at information set $\mathcal{I}_j \in \mathcal{I}$ iff σ_i maximises $U_i(\mathcal{I}_j, (\sigma'_i, \boldsymbol{\sigma}_{-i}), \boldsymbol{\beta})$ among all possible behaviour strategies σ'_i:
$$U_i(\mathcal{I}_j, (\sigma_i, \boldsymbol{\sigma}_{-i}), \boldsymbol{\beta}) = \max_{\sigma'_i \in \Sigma_i} U_i(\mathcal{I}_j, (\sigma'_i, \boldsymbol{\sigma}_{-i}), \boldsymbol{\beta})$$
2. Assessment $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ is **sequentially rational** iff for all players $i \in P$ and all information sets $\mathcal{I}_j \in p^{-1}(i)$, strategy σ_i is a best response to $\boldsymbol{\sigma}_{-i}$ w.r.t. $\boldsymbol{\beta}$ at $\mathcal{I}_j$.

Example (Simplified Poker)

- In $(\boldsymbol{\sigma}', \boldsymbol{\beta}')$ seen earlier, σ'_{Bob} is a best response to σ'_{Ann} at $\mathcal{I}_{\mathrm{B2}}$, because any $\sigma''_{\mathrm{Bob}}(\mathcal{I}_{\mathrm{B2}}) = \{\texttt{fold} \mapsto (1-q), \texttt{call} \mapsto q\}$ would likewise achieve a payoff of $U_{\mathrm{Bob}}(\mathcal{I}_{\mathrm{B2}}, (\sigma'_{\mathrm{Ann}}, \sigma''_{\mathrm{Bob}}), \beta') = \frac{1}{4} \cdot (-1 + q + 2q) + \frac{3}{4} \cdot (-1 + q - 2q) = \frac{-1+3q-3-3q}{4} = -1$.
- In contrast, σ'_{Ann} is not a best response to σ'_{Bob} at $\mathcal{I}_{\mathrm{A1}}$ as we shall see.

Consistent Beliefs: Example

In $(\boldsymbol{\sigma}', \boldsymbol{\beta}')$ seen earlier, we had

$$\sigma'_{\text{Ann}}(\mathcal{I}_{\text{A1}}) = \left\{ \texttt{check} \mapsto \frac{1}{2}, \texttt{raise} \mapsto \frac{1}{2} \right\}, \text{ and}$$

$$\beta'_{\text{Bob}}(\mathcal{I}_{\text{B2}}) = \left\{ [\texttt{deal123}, \texttt{raise}] \mapsto \frac{1}{4}, [\texttt{deal321}, \texttt{raise}] \mapsto \frac{3}{4} \right\}$$

However, Bob's beliefs about $\mathcal{I}_{\text{B2}}$ seem inadequate, as

$$P([\texttt{deal123}, \texttt{raise}] \,|\, \boldsymbol{\sigma}') = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12} \text{ and}$$

$$P([\texttt{deal321}, \texttt{raise}] \,|\, \boldsymbol{\sigma}') = \frac{1}{6} \cdot 1 = \frac{1}{6} = 2 \cdot P([\texttt{deal123}, \texttt{raise}] \,|\, \boldsymbol{\sigma}')$$

A more realistic likelihood estimate of the situation given by $\boldsymbol{\sigma}'$ would be

$$\beta''_{\text{Bob}}(\mathcal{I}_{\text{B2}}) = \left\{ [\texttt{deal123}, \texttt{raise}] \mapsto \frac{1}{3}, [\texttt{deal321}, \texttt{raise}] \mapsto \frac{2}{3} \right\}$$

Weakly Consistent Beliefs: Definition

Definition

Let G be an extensive-form game and $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ be an assessment for G.

Assessment $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ is **weakly consistent** iff for all information sets $\mathcal{I}_j \in \mathcal{I}$ and for all histories $h \in \mathcal{I}_j$, we have:

$$\beta_{p(\mathcal{I}_j)}(h \mid \mathcal{I}_j) = \frac{P(h \mid \boldsymbol{\sigma})}{\sum_{h \in \mathcal{I}_j} P(h \mid \boldsymbol{\sigma})} = \frac{P(h \mid \boldsymbol{\sigma})}{P(\mathcal{I}_j \mid \boldsymbol{\sigma})} \qquad \text{whenever } P(\mathcal{I}_j \mid \boldsymbol{\sigma}) > 0$$

Example (Simplified Poker)

The assessment $(\boldsymbol{\sigma}', \boldsymbol{\beta}')$ seen earlier is not weakly consistent.

Observation

Given a profile $\boldsymbol{\sigma}$ of behaviour strategies, we can use the definition above to construct a belief system $\boldsymbol{\beta}$ that is weakly consistent.

Weak Sequential Equilibria

Definition

Let G be an extensive-form game.

An assessment $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ for G is a **weak sequential equilibrium** iff it is both sequentially rational and weakly consistent.

Theorem (Kreps and Wilson, 1982)

Every extensive-form game with perfect recall and a finite set H of histories has a weak sequential equilibrium.

Recall: Perfect recall means that players know their own previous moves.

Example

Simplified Poker has perfect recall and is finite, therefore has a weak sequential equilibrium.

Some Special Cases

Theorem

Let G be a sequential game with perfect information and G' its associated extensive-form game (using singleton information sets).

The subgame-perfect equilibria of G and the weak sequential equilibria of G' are in one-to-one correspondence.

Theorem

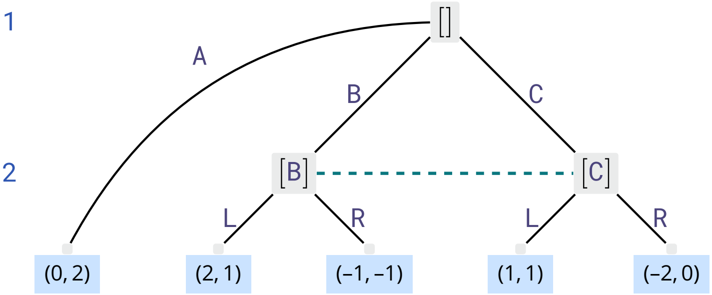
Let G be a strategic (normal-form) game (with simultaneous moves) and G' be its associated extensive-form game (using sequentialised moves and move hiding).

The mixed Nash equilibria of G and the weak sequential equilibria of G' are in one-to-one correspondence.

In both cases, we add (whenever necessary) a weakly consistent belief system.

Weak Sequential Equilibrium: (Non-)Example

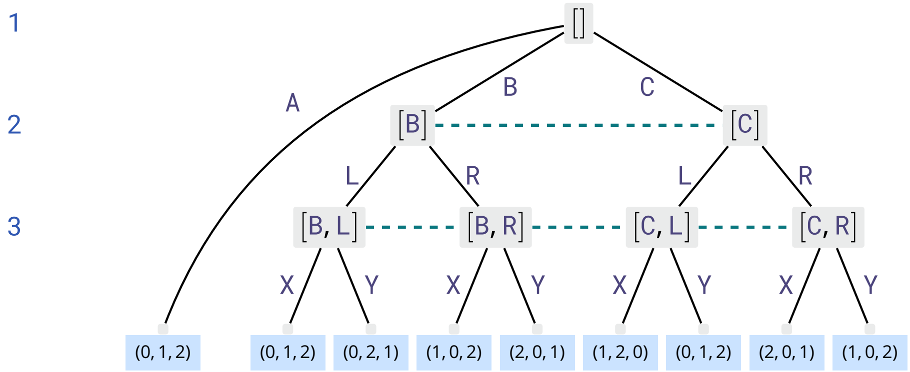
Reconsider game G_4:



- One (sequentially dubious) equilibrium of the normal form is $\mathbf{s} = (\mathsf{A}, \mathsf{R})$.
- One assessment resulting from this is $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ with $\boldsymbol{\sigma} \triangleq \mathbf{s}$, and $\beta_2(\mathcal{I}_2) = \{[\mathsf{B}] \mapsto x, [\mathsf{C}] \mapsto 1 - x\}$ being weakly consistent for any $x \in [0, 1]$.
- We have $U_2(\mathsf{L}, \boldsymbol{\beta}) = 1 > -x = U_2(\mathsf{R}, \boldsymbol{\beta})$ and $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ is not sequentially rational.

Why "Weak" Sequential Equilibria?

Consider this slight variation of G_4, the game G_5:



- We have that $\mathbf{s} = (\mathsf{A}, \mathsf{L}, \mathsf{Y})$ is a pure equilibrium of the normal form.
- For $(\mathbf{s}, \boldsymbol{\beta})$ let $\beta_2(\mathcal{I}_2) = \left\{[\mathsf{B}] \mapsto \frac{2}{3}, [\mathsf{C}] \mapsto \frac{1}{3}\right\}$ and $\beta_3(\mathcal{I}_3) = \left\{[\mathsf{B}, \mathsf{L}] \mapsto \frac{1}{3}, [\mathsf{C}, \mathsf{L}] \mapsto \frac{2}{3}, \ldots\right\}$.
- Player 3's beliefs about $\mathcal{I}_3$ are weakly consistent, but not justified.

Sequential Equilibria

Definition

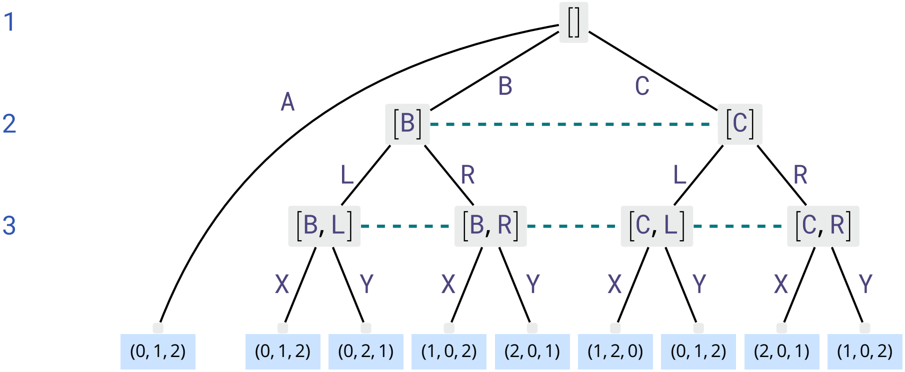
Let G be an extensive-form game with players $P = \{1, \ldots, n\}$.

1. Denote by $\boldsymbol{\Sigma}^{+}$ the set of all behaviour strategy profiles $(\sigma_1, \ldots, \sigma_n)$ such that for all $1 \le i \le n$, for all $\mathcal{I}_j \in \mathcal{I}$ and for all $m \in M_i(\mathcal{I}_j)$: $\quad \sigma_i(m \,|\, \mathcal{I}_j) > 0$.
2. For a given behaviour strategy profile $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}^{+}$, denote by $\boldsymbol{\beta}(\boldsymbol{\sigma})$ the unique (weakly consistent) belief system obtained from $\boldsymbol{\sigma}$ using Bayes' rule.
3. An assessment $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ is **consistent** iff there exists a sequence $\left(\boldsymbol{\sigma}^{(\ell)}, \boldsymbol{\beta}^{(\ell)}\right)_{\ell \in \mathbb{N}}$ with $\boldsymbol{\sigma}^{(\ell)} \in \boldsymbol{\Sigma}^{+}$ and $\boldsymbol{\beta}^{(\ell)} = \boldsymbol{\beta}(\boldsymbol{\sigma}^{(\ell)})$ for all $\ell \in \mathbb{N}$ such that

$$(\boldsymbol{\sigma}, \boldsymbol{\beta}) = \left(\lim_{\ell \to \infty} \boldsymbol{\sigma}^{(\ell)}, \lim_{\ell \to \infty} \boldsymbol{\beta}^{(\ell)} \right)$$

4. An assessment $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ is a **sequential equilibrium** iff $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ is sequentially rational and consistent.

Sequential Equilibria: Example



For $\mathbf{s} = (C, L, Y)$ and $\boldsymbol{\sigma} \triangleq \mathbf{s}$, we have that $(\boldsymbol{\sigma}, \boldsymbol{\beta}(\boldsymbol{\sigma}))$ is a sequential equilibrium:

$$\left(\sigma_1^{(\ell)}\right)_{\ell \in \mathbb{N}} = \left(\left\{ A \mapsto \frac{1}{10^{\ell+1}}, B \mapsto \frac{1}{10^{\ell+1}}, C \mapsto 1 - \frac{2}{10^{\ell+1}} \right\} \right)_{\ell \in \mathbb{N}}$$

Similar constructions can be given for $\sigma_2^{(\ell)}$ and $\sigma_3^{(\ell)}$ for $\ell \in \mathbb{N}$.

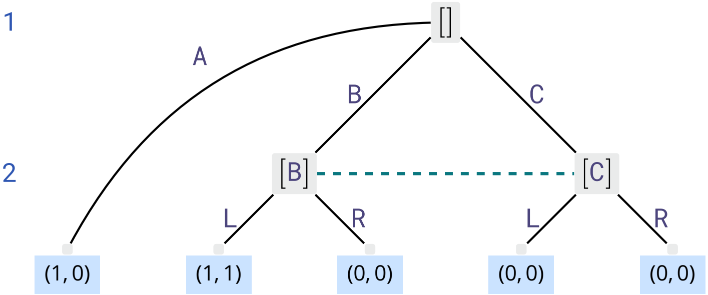
Perfect Sequential Equilibria

Definition

A sequential equilibrium $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ is **perfect** iff there exists a sequence $\left(\boldsymbol{\sigma}^{(\ell)}, \boldsymbol{\beta}^{(\ell)}\right)_{\ell \in \mathbb{N}}$ showing that $(\boldsymbol{\sigma}, \boldsymbol{\beta})$ is consistent where additionally:
for all players $i \in P$, strategy σ_i is a best response to $\boldsymbol{\sigma}_{-i}^{(\ell)}$ w.r.t. $\boldsymbol{\beta}^{(\ell)}$ for all $\ell \in \mathbb{N}$.

- Problematic cases typically arise with information sets not reached in equilibrium play.
- In $\boldsymbol{\Sigma}^+$, all players play all possible moves with positive probability.
- For $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}^+$, all information sets are reached and such unintended assessments are ruled out.
- Thus the sequence guarantees consistent beliefs in the limit, also at unreachable information sets.
- Requiring the limit to consist of best responses to the whole sequence guarantees resistance to small perturbations (“trembling hands”).

Perfect Sequential Equilibria: Non-Example

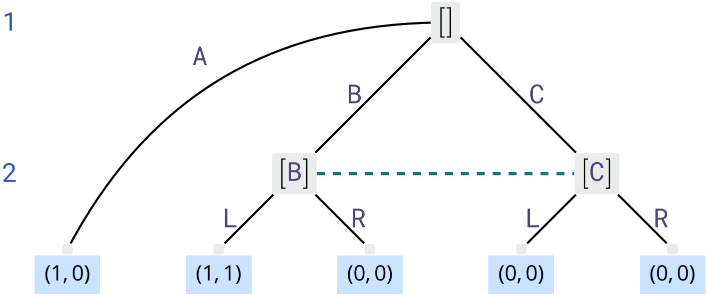


- (A, R) with $\beta_2([\mathsf{C}] \mid \mathcal{I}_2) = 1$ is a sequential equilibrium of G_6, but not perfect.
- Assume to the contrary that $\left(\boldsymbol{\sigma}^{(\ell)}, \boldsymbol{\beta}^{(\ell)}\right)_{\ell \in \mathbb{N}}$ witnesses (A, R) being perfect.
- Then for any $\ell \in \mathbb{N}$, we have $\sigma_1^{(\ell)}(\mathsf{B}) > 0$ and thus $\beta_2^{(\ell)}([\mathsf{B}] \mid \mathcal{I}_2) > 0$ whence

$$U_2(\mathcal{I}_2, (\mathsf{L}, \boldsymbol{\sigma}_{-i}^{(\ell)}), \boldsymbol{\beta}^{(\ell)}) = \beta_2^{(\ell)}([\mathsf{B}] \mid \mathcal{I}_2) > 0 = U_2(\mathcal{I}_2, (\mathsf{R}, \boldsymbol{\sigma}_{-i}^{(\ell)}), \boldsymbol{\beta}^{(\ell)})$$

and $\sigma_2 = \mathsf{R}$ is not a best response to $\boldsymbol{\sigma}_{-i}^{(\ell)}$ w.r.t. $\boldsymbol{\beta}^{(\ell)}$.

Perfect Sequential Equilibria: Example



- $(\sigma_1, \sigma_2) = (A, L) = ((1,0,0),(1,0))$ is a perfect sequential equilibrium:
- We reach σ_1 via $\left(1 - \frac{1}{2^\ell}, \frac{1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right)_{\ell\in\mathbb{N}}$ and σ_2 via $\left(1 - \frac{1}{2^{\ell+2}}, \frac{1}{2^{\ell+2}}\right)_{\ell\in\mathbb{N}}$.
- For any $\ell \in \mathbb{N}$, we have:

$$U_1(\sigma_1, \sigma_2^{(\ell)}) = U_1(A, \sigma_2^{(\ell)}) = 1 > U_1(B, \sigma_2^{(\ell)}) > U_1(C, \sigma_2^{(\ell)}) = 0,$$

$$U_2(\sigma_2, \sigma_1^{(\ell)}) = U_2(L, \sigma_1^{(\ell)}) = \beta_2^{(\ell)}([B] \mid \mathcal{I}_2) = \frac{1}{2} > U_2(R, \sigma_1^{(\ell)}) = 0.$$

Solving Simplified Poker

Solving Simplified Poker (1)

What happens in the two remaining cases?

Should Ann raise (i.e. bluff) if she has a 1?

Should Bob call (the bluff) if he has a 2?

- Denote by $\boldsymbol{\sigma}^* = (\sigma^*_{\text{Ann}}, \sigma^*_{\text{Bob}})$ the behaviour strategy profile where both players act optimally according to our previous analysis, and additionally
- Ann resolves to bluff (with a 1) with probability p, $\sigma^*_{\text{Ann}}(\texttt{raise} \mid \mathcal{I}_{\text{A1}}) = p$,
- Bob resolves to call (with a 2) with probability q, $\sigma^*_{\text{Bob}}(\texttt{call} \mid \mathcal{I}_{\text{B2}}) = q$.
- Denote by $\boldsymbol{\beta}^*$ the belief system that is consistent with $\boldsymbol{\sigma}^*$.
- We know $P([\texttt{deal123}] \mid \boldsymbol{\sigma}^*) = P([\texttt{deal132}] \mid \boldsymbol{\sigma}^*) = \frac{1}{6}$, so $P(\mathcal{I}_{\text{A1}} \mid \boldsymbol{\sigma}^*) = \frac{1}{3}$ and
- $\beta^*_{\text{Ann}}([\texttt{deal123}] \mid \mathcal{I}_{\text{A1}}) = \beta^*_{\text{Ann}}([\texttt{deal132}] \mid \mathcal{I}_{\text{A1}}) = \frac{1}{2}$.

1. How should Ann choose the value of p?
2. How should Bob choose the value of q?

Solving Simplified Poker (2)

$$\begin{aligned}P(\mathcal{I}_{B2} \mid \boldsymbol{\sigma}^*) &= P([\texttt{deal123}, \texttt{raise}] \mid \boldsymbol{\sigma}^*) + P([\texttt{deal321}, \texttt{raise}] \mid \boldsymbol{\sigma}^*)\\ &= P([\texttt{deal123}] \mid \boldsymbol{\sigma}^*) \cdot \sigma^*_{\text{Ann}}(\texttt{raise} \mid \mathcal{I}_{A1}) + P([\texttt{deal321}] \mid \boldsymbol{\sigma}^*) \cdot \sigma^*_{\text{Ann}}(\texttt{raise} \mid \mathcal{I}_{A3})\\ &= \frac{1}{6} \cdot p + \frac{1}{6} \cdot 1\end{aligned}$$

Therefore,

$$P([\texttt{deal123}, \texttt{raise}] \mid \mathcal{I}_{B2}, \boldsymbol{\sigma}^*) = \frac{P([\texttt{deal123}, \texttt{raise}] \mid \boldsymbol{\sigma}^*)}{P(\mathcal{I}_{B2} \mid \boldsymbol{\sigma}^*)} = \frac{\frac{p}{6}}{\frac{p}{6} + \frac{1}{6}} = \frac{p}{p+1}$$

$$P([\texttt{deal321}, \texttt{raise}] \mid \mathcal{I}_{B2}, \boldsymbol{\sigma}^*) = \frac{P([\texttt{deal321}, \texttt{raise}] \mid \boldsymbol{\sigma}^*)}{P(\mathcal{I}_{B2} \mid \boldsymbol{\sigma}^*)} = \frac{\frac{1}{6}}{\frac{p}{6} + \frac{1}{6}} = \frac{1}{p+1}$$

Ann's goal is to make Bob indifferent between his two moves in $\mathcal{I}_{B2}$, that is:

$$U_{\text{Bob}}(\texttt{fold}, \mathcal{I}_{B2}, \boldsymbol{\sigma}^*) = U_{\text{Bob}}(\texttt{call}, \mathcal{I}_{B2}, \boldsymbol{\sigma}^*)$$

Solving Simplified Poker (3)

We have the below payoff when Bob plays `fold` at $\mathcal{I}_{B2}$ with probability 1:

$$U_{Bob}(\texttt{fold}, \mathcal{I}_{B2}, \boldsymbol{\sigma}^*) = -1$$

and likewise, if Bob plays a pure `call` at $\mathcal{I}_{B2}$:

$$\begin{aligned} U_{Bob}(\texttt{call}, \mathcal{I}_{B2}, \boldsymbol{\sigma}^*) &= P([\texttt{deal123}, \texttt{raise}] \mid \mathcal{I}_{B2}, \boldsymbol{\sigma}^*) \cdot u_{Bob}([\texttt{deal123}, \texttt{raise}, \texttt{call}]) + \\ &\quad\; P([\texttt{deal321}, \texttt{raise}] \mid \mathcal{I}_{B2}, \boldsymbol{\sigma}^*) \cdot u_{Bob}([\texttt{deal321}, \texttt{raise}, \texttt{call}]) \\ &= \frac{p}{p+1} \cdot 2 + \frac{1}{p+1} \cdot (-2) = \frac{2p-2}{p+1} \end{aligned}$$

So overall, Ann's goal is to choose p such that

$$-1 = \frac{2p-2}{p+1} \qquad \text{whence we obtain} \qquad p = \frac{1}{3}.$$

Solving Simplified Poker (4)

It remains to calculate $q = \sigma^*_{\text{Bob}}(\texttt{call} \mid \mathcal{I}_{\text{B2}})$.

Again, Bob's goal is to make Ann indifferent between her two moves in $\mathcal{I}_{\text{A1}}$:

$$U_{\text{Ann}}(\texttt{check}, \mathcal{I}_{\text{A1}}, \boldsymbol{\sigma}^*) = U_{\text{Ann}}(\texttt{raise}, \mathcal{I}_{\text{A1}}, \boldsymbol{\sigma}^*)$$

For the left-hand side, we obtain the expected utility of a pure check at $\mathcal{I}_{\text{A1}}$:

$$U_{\text{Ann}}(\texttt{check}, \mathcal{I}_{\text{A1}}, \boldsymbol{\sigma}^*) = -1$$

Solving Simplified Poker (5)

For the right-hand side, we get the expected utility of a pure `raise` at $\mathcal{I}_{A1}$:

$$
\begin{aligned}
& U_{\text{Ann}}(\texttt{raise}, \mathcal{I}_{A1}, \boldsymbol{\sigma}^*) \\
= {} & P([\texttt{deal123}] \mid \mathcal{I}_{A1}, \boldsymbol{\sigma}^*) \cdot \sigma_{\text{Bob}}(\texttt{fold} \mid \mathcal{I}_{B2}) \cdot u_{\text{Ann}}([\texttt{deal123}, \texttt{raise}, \texttt{fold}]) + \\
& P([\texttt{deal123}] \mid \mathcal{I}_{A1}, \boldsymbol{\sigma}^*) \cdot \sigma_{\text{Bob}}(\texttt{call} \mid \mathcal{I}_{B2}) \cdot u_{\text{Ann}}([\texttt{deal123}, \texttt{raise}, \texttt{call}]) + \\
& P([\texttt{deal132}] \mid \mathcal{I}_{A1}, \boldsymbol{\sigma}^*) \cdot \sigma_{\text{Bob}}(\texttt{fold} \mid \mathcal{I}_{B3}) \cdot u_{\text{Ann}}([\texttt{deal132}, \texttt{raise}, \texttt{fold}]) + \\
& P([\texttt{deal132}] \mid \mathcal{I}_{A1}, \boldsymbol{\sigma}^*) \cdot \sigma_{\text{Bob}}(\texttt{call} \mid \mathcal{I}_{B3}) \cdot u_{\text{Ann}}([\texttt{deal132}, \texttt{raise}, \texttt{call}]) \\
= {} & \frac{1}{2} \cdot (1 - q) \cdot 1 + \frac{1}{2} \cdot q \cdot (-2) + \frac{1}{2} \cdot 0 \cdot 1 + \frac{1}{2} \cdot 1 \cdot (-2) = \frac{1}{2} \cdot (1 - q - 2q - 2)
\end{aligned}
$$

Overall, Bob's goal is thus to choose q such that

$$-1 = \frac{-3q - 1}{2} \qquad \text{whence we obtain} \qquad q = \frac{1}{3}.$$

Solving Simplified Poker: Takeaways

- Bluffing can be part of a rational strategy (playing against rational opponents):
  - Ann bluffs a third of the times she has her worst possible hand,
  - which is justified because Bob will call that raise only a third of the times.
- The expected value of the game for the obtained $\boldsymbol{\sigma}^*$ is

$$U_{\text{Ann}}(\boldsymbol{\sigma}^*) = \frac{p - 3pq + q}{6} = \frac{1}{18} = -U_{\text{Bob}}(\boldsymbol{\sigma}^*)$$

  so Ann has an advantage. Thus players switch roles after each round.
- If Ann deviates from $\boldsymbol{\sigma}^*$, then Bob will best-respond (punish) by adapting q:
  - for $p > \frac{1}{3}$ setting $q = 1$, and
  - for $p < \frac{1}{3}$ setting $q = 0$.

Solving (heads-up limit Texas hold'em) Poker

Bowling et al. [2015] consider heads-up limit hold'em poker to be "essentially weakly solved":

- There are $3.16 \cdot 10^{17}$ possible states, and $3.19 \cdot 10^{14}$ decision points.
- They used an algorithm called counterfactual regret minimisation$^+$ (CFR$^+$):
  - Uses self-play and in hindsight, computes regret (utility difference to best decision) of taken moves.
  - Obtains successive approximations to a Nash equilibrium.
  - Took 900 core-years of computation, on 200 nodes of 24 cores each.
  - Solution quality can be assessed via so-called exploitability:
    Expected loss of the computed strategy against the worst-case opponent.
- Essentially solved: Lifetime of play ($70y \cdot 365d \cdot 12h \cdot 200$ games) cannot statistically differentiate the game from being solved (at 95% confidence).
- Game-theoretic value is between 87.7 and 89.7 *mbb/g* (milli-big-blinds per game) for the dealer (the player moving first).

Conclusion

Summary

- A **belief system** assigns probabilities to histories in information sets.
- An **assessment** is a pair (behaviour strategy profile, belief system).
- A **sequentially rational** assessment plays best responses “everywhere”.
- An assessment is **weakly consistent** whenever the belief system’s probabilities match what is expected from everyone playing according to the behaviour strategy profile.
- An assessment is **consistent** iff it is the limit of a (justifying) sequence of assessments with all-positive-probability strategies.
- An assessment is a **(weak) sequential equilibrium** iff it is both sequentially rational and (weakly) consistent.
- A sequential equilibrium is **perfect** iff every player plays a best response against every opponent profile in the equilibrium’s justifying sequence.