TECHNISCHE UNIVERSITÄT DRESDEN

Computational Logic ∴ Group

Hannes Strass

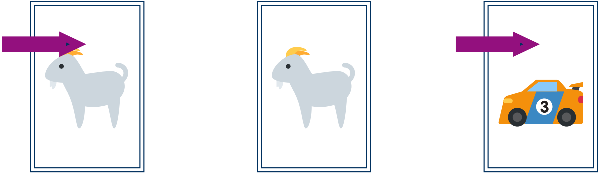
Faculty of Computer Science, Institute of Artificial Intelligence, Computational Logic Group

Games with Missing Information: Solving

Lecture 7, 3rd Jun 2024 // Algorithmic Game Theory, SS 2024

Previously . . .

- In **complete information** games, players know the rules, possible outcomes and each other's preferences over outcomes.
- In **perfect information** games, moves are sequential and all players know all previous moves.
- In **extensive-form** games, information is not necessarily complete or perfect.
- Uncertainty of players (due to missing information) can be modelled by **information sets** and **chance nodes** (moves by `Nature`).
- **Bayes' Theorem** shows how to compute with conditional probabilities.
- The **law of total probability** relates marginal to conditional probabilities.



Overview

Example: Simplified Poker

Behaviour Strategies and Belief Systems

Weak Sequential Equilibria

Solving Simplified Poker

Example: Simplified Poker

Simplified Poker: Game Description

Binmore's Simplified Poker

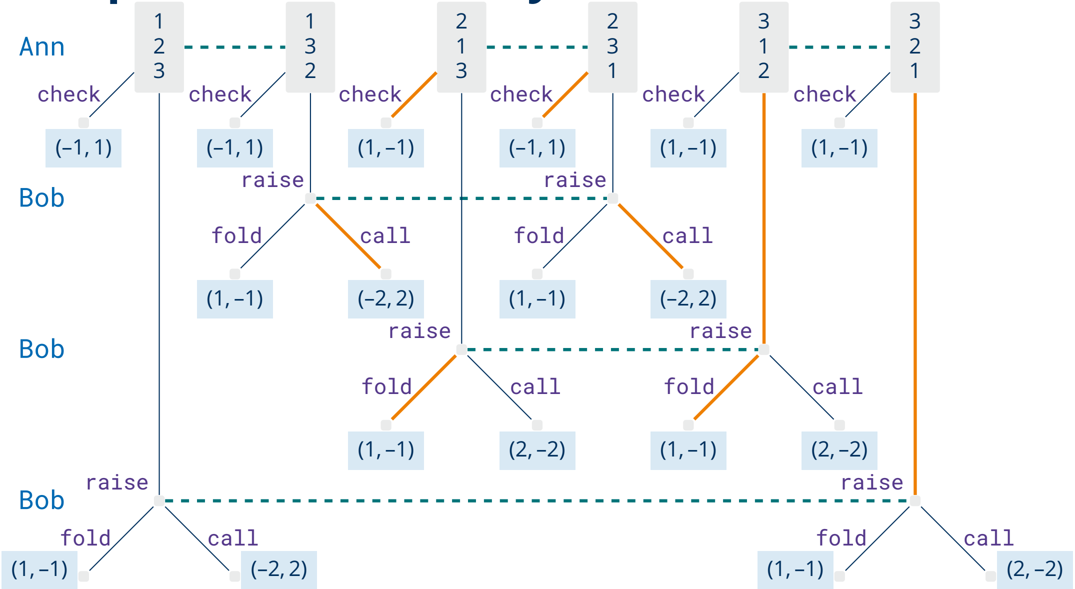
- Two players, Ann and Bob, each put \$1 into a jackpot.
- They then draw one card from a deck of three cards: $\{1, 2, 3\}$.
- Ann can either `check` (pass on), or `raise` (put another \$1 into the jackpot).
- Next, Bob responds:
  - If Ann has `check`ed, then Bob must `call`, that is, a showdown happens: 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:

- $P = \{\mathtt{Ann}, \mathtt{Bob}, \mathtt{Nature}\}$
- $\mathbf{M} = (M_{\mathtt{Ann}}, M_{\mathtt{Bob}}, M_{\mathtt{Nature}})$ with
  - $M_{\mathtt{Ann}} = \{\mathtt{check}, \mathtt{raise}\}$,
  - $M_{\mathtt{Bob}} = \{\mathtt{fold}, \mathtt{call}\}$,
  - $M_{\mathtt{Nature}} = \{\mathtt{deal123}, \mathtt{deal132}, \mathtt{deal213}, \mathtt{deal231}, \mathtt{deal312}, \mathtt{deal321}\}$.
- $\mathcal{I} = \{\mathcal{I}_{A1}, \mathcal{I}_{A2}, \mathcal{I}_{A3}, \mathcal{I}_{B1}, \mathcal{I}_{B2}, \mathcal{I}_{B3}\}$ with
  - $\mathcal{I}_{A1} = \{[\mathtt{deal123}], [\mathtt{deal132}]\}$,
    $\mathcal{I}_{A2} = \{[\mathtt{deal213}], [\mathtt{deal231}]\}$,
    $\mathcal{I}_{A3} = \{[\mathtt{deal312}], [\mathtt{deal321}]\}$ with $p(\mathcal{I}_{A1}) = p(\mathcal{I}_{A2}) = p(\mathcal{I}_{A3}) = \mathtt{Ann}$,
  - $\mathcal{I}_{B1} = \{[\mathtt{deal213}, \mathtt{raise}], [\mathtt{deal312}, \mathtt{raise}]\}$,
    $\mathcal{I}_{B2} = \{[\mathtt{deal123}, \mathtt{raise}], [\mathtt{deal321}, \mathtt{raise}]\}$,
    $\mathcal{I}_{B3} = \{[\mathtt{deal132}, \mathtt{raise}], [\mathtt{deal231}, \mathtt{raise}]\}$ with $p(\mathcal{I}_{B1}) = p(\mathcal{I}_{B2}) = p(\mathcal{I}_{B3}) = \mathtt{Bob}$.
- $\mathbf{u} = (u_{\mathtt{Ann}}, u_{\mathtt{Bob}})$ 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?

Behaviour Strategies and Belief Systems

Behaviour Strategies (1)

Definition

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

1. A **pure strategy** for player $i \in P$ is a function s_i that assigns a possible move to each of player i's information sets.
2. A **behaviour strategy** for player $i \in P$ is a function π_i that assigns a probability distribution over possible moves to each of player i's information sets.

- $s_i(\mathcal{I}_j)$ denotes the move taken by player i at information set $\mathcal{I}_j \in \mathcal{I}$.
- $\pi_i(\mathcal{I}_j)(m_k)$ is the probability that player i will make move m_k at information set $\mathcal{I}_j$. For readability, we will write this as $\pi_i(m_k \mid \mathcal{I}_j)$.
- As usual, a pure strategy s_i with $s_i(\mathcal{I}_j) = m_k$ can be seen as a behaviour strategy π_i with $\pi_i(m_k \mid \mathcal{I}_j) = 1$ and $\pi_i(m_\ell \mid \mathcal{I}_j) = 0$ for $m_\ell \in M_i$, $\ell \neq k$.

Behaviour Strategies (2)

Example (Simplified Poker)

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

A behaviour strategy profile $\boldsymbol{\pi}$ induces expected utilities for all players:

$$U_i(\boldsymbol{\pi}) := \sum_{z \in Z} P(z \mid \boldsymbol{\pi}) \cdot u_i(z)$$

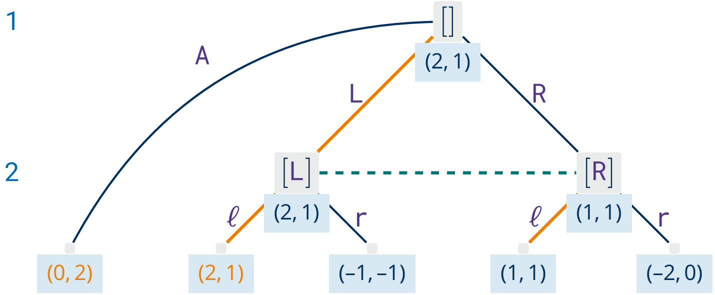
where $P(h \mid \boldsymbol{\pi})$ is the probability that history h is reached whenever play happens according to profile $\boldsymbol{\pi}$: inductively, define $P([] \mid \boldsymbol{\pi}) := 1$ and

$$P([h; m] \mid \boldsymbol{\pi}) := \pi_{p(\mathcal{I}_h)}(m \mid \mathcal{I}_h) \cdot P(h \mid \boldsymbol{\pi})$$

- where $\mathcal{I}_h \in \mathcal{I}$ is the unique information set with $h \in \mathcal{I}_h$,
- and $\pi_{\mathtt{Nature}}$ 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)
L	(2, 1)	(–1, –1)
R	(1, 1)	(–2, 0)

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

⇝ 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 $(\mathtt{A}, \mathtt{r})$ is "subgame perfect".
- In $(\mathtt{A}, \mathtt{r})$, information set $\{[\mathtt{L}], [\mathtt{R}]\}$ is reached with probability zero.
- To define playing best responses "everywhere": What is the expected payoff from information set $\{[\mathtt{L}], [\mathtt{R}]\}$ when play happens as in $(\mathtt{A}, \mathtt{r})$?

$\rightsquigarrow$ 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 $(\boldsymbol{\pi}, \boldsymbol{\beta})$ consisting of a profile $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_n)$ of behaviour strategies and a belief system $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$.

Example (Simplified Poker)

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

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

Expected Utility for Assessments

Definition

Let G be an extensive-form game and $(\boldsymbol{\pi}, \boldsymbol{\beta})$ be an assessment of G.
The **expected utility** for player i at information set $\mathcal{I}_j$ according to $(\boldsymbol{\pi}, \boldsymbol{\beta})$ is

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

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

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

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

Weak Sequential Equilibria

Best Responses and Sequential Rationality

Definition

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

1. Player i's strategy π_i is a **best response** to $\boldsymbol{\pi}_{-i}$ at information set $\mathcal{I}_j \in \mathcal{I}$ iff π_i maximises $U_i(\mathcal{I}_j, (\boldsymbol{\pi}_{-i}, \pi_i'), \boldsymbol{\beta})$ among all possible behaviour strategies π_i':
$$U_i(\mathcal{I}_j, (\boldsymbol{\pi}_{-i}, \pi_i, \boldsymbol{\beta})) = \max_{\pi_i' \in \Pi_i} U_i(\mathcal{I}_j, (\boldsymbol{\pi}_{-i}, \pi_i', \boldsymbol{\beta}))$$
2. Assessment $(\boldsymbol{\pi}, \boldsymbol{\beta})$ is **sequentially rational** iff for all players $i \in P$, strategy π_i is a best response at each information set $\mathcal{I}_j$ with $p(\mathcal{I}_j) \in \{i, \texttt{Nature}\}$.

Example (Simplified Poker)

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

Consistency of Beliefs: Example

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

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

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

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

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

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

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

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

Consistency of Beliefs: Definition

Definition

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

Assessment $(\boldsymbol{\pi}, \boldsymbol{\beta})$ satisfies **consistency of beliefs** iff for all information sets $\mathcal{I}_j \in \mathcal{I}$ and for all histories $h \in \mathcal{I}_j$, we have:

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

Example (Simplified Poker)

The assessment $(\boldsymbol{\pi}', \boldsymbol{\beta}')$ seen earlier does not satisfy consistency of beliefs.

Observation

Given a profile $\boldsymbol{\pi}$ of behaviour strategies, we can use the definition above to construct a belief system $\boldsymbol{\beta}$ that satisfies consistency of beliefs.

Weak Sequential Equilibria

Definition

Let G be an extensive-form game.

An assessment $(\boldsymbol{\pi}, \boldsymbol{\beta})$ for G is a **weak sequential equilibrium** iff it is both sequentially rational and satisfies consistency of beliefs.

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).

Every subgame-perfect equilibrium of G corresponds to a weak sequential equilibrium of G'.

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).

Every mixed Nash equilibrium of G corresponds to a weak sequential equilibrium of G'.

In both cases, we add a belief system satisfying consistency of beliefs.

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 $\boldsymbol{\pi}^* = (\pi^*_{\text{Ann}}, \pi^*_{\text{Bob}})$ 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, $\pi^*_{\text{Ann}}(\texttt{raise} \mid \mathcal{I}_{\text{A1}}) = p$,
- Bob resolves to call (with a 2) with probability q, $\pi^*_{\text{Bob}}(\texttt{call} \mid \mathcal{I}_{\text{B2}}) = q$.
- Denote by $\boldsymbol{\beta}^*$ the belief system that is consistent with $\boldsymbol{\pi}^*$.
- We know $P([\texttt{deal123}] \mid \boldsymbol{\pi}^*) = P([\texttt{deal132}] \mid \boldsymbol{\pi}^*) = \frac{1}{6}$, so $P(\mathcal{I}_{\text{A1}} \mid \boldsymbol{\pi}^*) = \frac{1}{3}$ 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)

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

Therefore,

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

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

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

$$
U_{\mathrm{Bob}}(\texttt{fold}, \mathcal{J}_{B2}, \boldsymbol{\pi}^*) = U_{\mathrm{Bob}}(\texttt{call}, \mathcal{J}_{B2}, \boldsymbol{\pi}^*)
$$

Solving Simplified Poker (3)

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

$$U_{Bob}(\texttt{fold}, \mathcal{I}_{B2}, \boldsymbol{\pi}^*) = P([\texttt{deal123}, \texttt{raise}] \mid \mathcal{I}_{B2}, \boldsymbol{\pi}^*) \cdot u_{Bob}([\texttt{deal123}, \texttt{raise}, \texttt{fold}]) + P([\texttt{deal321}, \texttt{raise}] \mid \mathcal{I}_{B2}, \boldsymbol{\pi}^*) \cdot u_{Bob}([\texttt{deal321}, \texttt{raise}, \texttt{fold}])$$

$$= \frac{p}{p+1} \cdot (-1) + \frac{1}{p+1} \cdot (-1) = -1$$

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

$$U_{Bob}(\texttt{call}, \mathcal{I}_{B2}, \boldsymbol{\pi}^*) = P([\texttt{deal123}, \texttt{raise}] \mid \mathcal{I}_{B2}, \boldsymbol{\pi}^*) \cdot u_{Bob}([\texttt{deal123}, \texttt{raise}, \texttt{call}]) + P([\texttt{deal321}, \texttt{raise}] \mid \mathcal{I}_{B2}, \boldsymbol{\pi}^*) \cdot u_{Bob}([\texttt{deal321}, \texttt{raise}, \texttt{call}])$$

$$= \frac{p}{p+1} \cdot 2 + \frac{1}{p+1} \cdot (-2) = \frac{2p-2}{p+1}$$

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 $q = \pi^*_{\text{Bob}}(\texttt{call} \mid \mathcal{I}_{\text{B2}})$.

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

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

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

$$\begin{aligned} & U_{\text{Ann}}(\texttt{check}, \mathcal{I}_{\text{A1}}, \boldsymbol{\pi}^*) \\ = \; & P([\texttt{deal123}] \mid \mathcal{I}_{\text{A1}}, \boldsymbol{\pi}^*) \cdot u_{\text{Ann}}([\texttt{deal123}, \texttt{check}]) + \\ & P([\texttt{deal132}] \mid \mathcal{I}_{\text{A1}}, \boldsymbol{\pi}^*) \cdot u_{\text{Ann}}([\texttt{deal132}, \texttt{check}]) \\ = \; & \frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot (-1) = -1 \end{aligned}$$

Solving Simplified Poker (5)

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

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

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 $\boldsymbol{\pi}^*$ is

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

  so Ann has an advantage. Thus players switch roles after each round.
- If Ann deviates from $\boldsymbol{\pi}^*$, 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 **behaviour strategy** assigns move probabilities to information sets.
- 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 satisfies **consistency of beliefs** whenever the belief system’s probabilities match what is expected from everyone playing according to the behaviour strategy profile.
- An assessment is a **weak sequential equilibrium** iff it is both sequentially rational and satisfies consistency of beliefs.
- Mixed Nash equilibria for normal-form games and subgame perfect equilibria for sequential perfect-information games are special cases of weak sequential equilibria for extensive-form games.