

CSCE 631 — CFR Worked Example: Kuhn Poker

L13's definitions, evaluated on one complete game

Alan Kuhnle

Lecture 13 companion Fall 2026

Texas A&M University

One game, every step of CFR

Kuhn poker setup

Reach probabilities

Counterfactual values

Counterfactual regret and regret matching

Three complete CFR iterations

Why the average strategy?

Summary

L13 connection: Follow the definitions in order, then run the simultaneous-update pseudocode.

The rules and the information available

- Two players ante 1 chip each. Deal one private card each, without replacement, from $J < Q < K$; each ordered deal has probability $1/6$.
- **Player 1** acts first: **check** or **bet 1**. After a check, **Player 2** checks for showdown or bets 1.
- Facing a bet, the other player **folds** or **calls 1**. There are no raises. The higher card wins at showdown.
- **Net chip profit:** $u_2 = -u_1$; ± 1 after check-check or a fold, and ± 2 after a call.

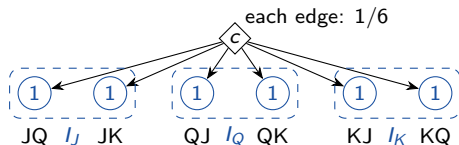
Write $x = \text{check}$, $b = \text{bet}$, $f = \text{fold}$, $c = \text{call}$ and $h = (d_1, d_2; s)$. Let $I_{d,s}^i \in \mathcal{I}_i$ identify own card d and public history s ; abbreviate $I_{d,\emptyset}^1$ by I_d .

Our running example

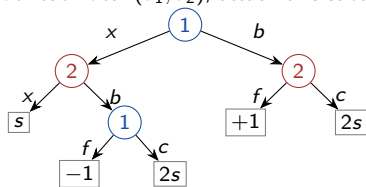
We will use this game as a running example to compute every quantity defined in L13.

L13 connection: Histories h , terminal histories z , information sets I , and utilities $u_i(z)$.

A compact tree: six deals, one shared betting topology



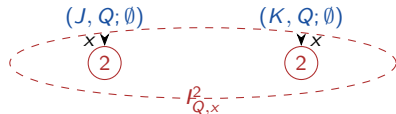
Under *each* deal (d_1, d_2) , attach this subtree:



Leaves: u_1 ; $u_2 = -u_1$. $s = \text{sgn}(d_1 - d_2)$.

Which nodes are indistinguishable? For each card d , these two histories form one information set:

Information set	Histories in it ($e \neq d$)
I_d	$(d, e; \emptyset)$
$I_{d,x}^1$	$(d, e; xb)$
$I_{d,x}^2$	$(e, d; x)$
$I_{d,b}^2$	$(e, d; b)$



P1: $3 + 3 = 6$ sets; P2: $3 + 3 = 6$ sets.

L13 connection: An information set groups histories, not cards or actions visible to both players.

A complete Nash equilibrium to use as a reference

Each entry gives the probability of the named action; its alternative has probability $1 - p$.

Player	Decision / named action	J	Q	K
P1	Opening: bet	$\frac{1}{3}$	0	1
P1	After check–bet: call	0	$\frac{2}{3}$	1
P2	After check: bet	$\frac{1}{3}$	0	1
P2	Facing opening bet: call	0	$\frac{1}{3}$	1

Verification: enumerating all $2^6 = 64$ pure plans for each player gives

$$u_1(\sigma^*) = -\frac{1}{18} = \max_{\sigma'_1} u_1(\sigma'_1, \sigma_2^*), \quad u_2(\sigma^*) = \frac{1}{18} = \max_{\sigma'_2} u_2(\sigma_1^*, \sigma'_2).$$

This is one equilibrium. More generally, P1 can open-bet with J with probability $\alpha \in [0, 1/3]$, bet K with probability 3α , and call with Q after check–bet with probability $\alpha + 1/3$. The table chooses $\alpha = 1/3$; P2's displayed strategy stays fixed.

L13 connection: Nash equilibrium means neither player gains from a whole-strategy deviation.

Uniform play: multiply only actions before the node

Start at $t = 1$: $\sigma_i^1(a \mid I) = 1/2$ at all information sets. For $h = (J, Q; s)$, compute the reach *before* its next action:

$$\pi^\sigma(h) = \pi_1^\sigma(h)\pi_2^\sigma(h)\pi_c(h), \quad \pi_{-i}^\sigma(h) = \pi_c(h) \prod_{j \neq i} \pi_j^\sigma(h).$$

Node / acting player	π_1	π_2	π_c	π^σ	π_{-i}^σ
$(J, Q; \emptyset) / P1$	1	1	$\frac{1}{6}$	$1 \cdot 1 \cdot \frac{1}{6} = \frac{1}{6}$	$\frac{1}{6}$
$(J, Q; x) / P2$	$\frac{1}{2}$	1	$\frac{1}{6}$	$\frac{1}{2} \cdot 1 \cdot \frac{1}{6} = \frac{1}{12}$	$\frac{1}{12}$
$(J, Q; xb) / P1$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{24}$	$\frac{1}{12}$

Counterfactual reach

This is the probability of reaching h if player i tries to get there — everyone else's contribution only, including chance.

L13 connection: Counterfactual reach excludes only the acting player's earlier action probabilities.

One information set, two counterfactual weights

Player 2 holds Q after a check:

$$I_{Q,x}^2 = \{(J, Q; x), (K, Q; x)\}, \quad \pi_{-2}^\sigma(d, Q; x) = \frac{1}{6} \sigma_1(x \mid I_d).$$

Node	Uniform σ^1		Preview: computed σ^3	
	P1 checks	$\pi_{-2}^{\sigma^1}$	P1 checks	$\pi_{-2}^{\sigma^3}$
$(J, Q; x)$	$\frac{1}{2}$	$\frac{1}{12}$	$\frac{5}{8}$	$\frac{5}{48}$
$(K, Q; x)$	$\frac{1}{2}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{24}$

Under uniform play the weights are **equal**. Once P1 checks at different rates with different cards, the weights can differ within the same information set.

Do not normalize these weights to sum to one. L13's counterfactual value uses the raw chance-and-opponent reach probabilities.

L13 connection: $\pi_{-i}^\sigma(h)$ belongs to a history h ; one I may contain unequal weights.

P1 holds J: enumerate every terminal continuation

$I_J = \{(J, Q; \emptyset), (J, K; \emptyset)\}$. At either node, $\pi_{-1}^{\sigma^1}(h) = 1/6$ and the continuation table is identical:

Terminal	Actions	$\pi^{\sigma^1}(h, z)$	$u_1(z)$
xx	check, check	$\frac{1}{4}$	-1
xbf	check, bet, fold	$\frac{1}{8}$	-1
xbc	check, bet, call	$\frac{1}{8}$	-2
bf	bet, fold	$\frac{1}{4}$	1
bc	bet, call	$\frac{1}{4}$	-2

$$\begin{aligned} v_1^{\sigma^1}(I_J) &= \sum_{h \in I_J} \pi_{-1}^{\sigma^1}(h) \sum_{z \in Z_h} \pi^{\sigma^1}(h, z) u_1(z) \\ &= \underbrace{\left(\frac{1}{6} + \frac{1}{6}\right)}_{\text{two private deals}} \underbrace{\left[\frac{1}{4}(-1) + \frac{1}{8}(-1) + \frac{1}{8}(-2) + \frac{1}{4}(1) + \frac{1}{4}(-2)\right]}_{-\frac{7}{8}} = \boxed{-\frac{7}{24}}. \end{aligned}$$

L13 connection: This is exactly $v_i^{\sigma}(I)$ in L13's "Counterfactual Value" definition.

Force check or bet: remove the first action's probability

L13 defines the action value using continuation from $h \cdot a$:

$$v_i^\sigma(I, a) = \sum_{h \in I} \pi_{-i}^\sigma(h) \sum_{z \in Z_{h \cdot a}} \pi^\sigma(h \cdot a, z) u_i(z).$$

For a forced check with J , the continuation probabilities are $1/2$ (check), $1/4$ (bet-fold), $1/4$ (bet-call):

$$v_1^{\sigma^1}(I_J, \text{check}) = \left(\frac{1}{6} + \frac{1}{6}\right) \left[\frac{1}{2}(-1) + \frac{1}{4}(-1) + \frac{1}{4}(-2)\right] = \boxed{-\frac{5}{12}}.$$

For a forced bet, P2 folds or calls with probability $1/2$ each:

$$v_1^{\sigma^1}(I_J, \text{bet}) = \left(\frac{1}{6} + \frac{1}{6}\right) \left[\frac{1}{2}(1) + \frac{1}{2}(-2)\right] = \boxed{-\frac{1}{6}}.$$

Check the decomposition:

$$v_1^{\sigma^1}(I_J) = \frac{1}{2} \left(-\frac{5}{12}\right) + \frac{1}{2} \left(-\frac{1}{6}\right) = -\frac{7}{24}.$$

L13 connection: The counterfactual value at an information set mixes the forced-action values.

P2 holds Q facing a bet: use P2's payoff

$I_{Q,b}^2 = \{(J, Q; b), (K, Q; b)\}$. Each node has $\pi_{-2}^{\sigma^1} = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$.

Hidden P1 card	CF weight	u_2 if fold	u_2 if call
J	$\frac{1}{12}$	-1	+2
K	$\frac{1}{12}$	-1	-2

After either forced action the game ends, so its continuation probability is 1.

$$v_2^{\sigma^1}(I_{Q,b}^2, \text{fold}) = \frac{1}{12}(-1) + \frac{1}{12}(-1) = \boxed{-\frac{1}{6}},$$

$$v_2^{\sigma^1}(I_{Q,b}^2, \text{call}) = \frac{1}{12}(2) + \frac{1}{12}(-2) = \boxed{0},$$

$$v_2^{\sigma^1}(I_{Q,b}^2) = \frac{1}{2} \left(-\frac{1}{6} \right) + \frac{1}{2}(0) = \boxed{-\frac{1}{12}}.$$

L13 connection: The same action-value definition applies to $i = 2$; use $u_2 = -u_1$.

Think about it: P1 holds K

Before the next slide

Try computing $v_1^{\sigma^1}(I_K)$ yourself.

$I_K = \{(K, J; \emptyset), (K, Q; \emptyset)\}$. Keep every action uniform.

1. Compute the value after forcing **check**.
2. Compute the value after forcing **bet**.
3. Mix those two values with probabilities $1/2, 1/2$.

Watch the fold: even a king loses its ante if P1 folds!

L13 connection: Practice the two nested sums in the counterfactual action-value definition.

P1 holds K: the answer

At either private-deal node, checking leads to payoff +1 after xx , -1 after xbf , and $+2$ after xbc .

$$v_1^{\sigma^1}(I_K, \text{check}) = \left(\frac{1}{6} + \frac{1}{6}\right) \left[\frac{1}{2}(1) + \frac{1}{4}(-1) + \frac{1}{4}(2)\right] = \boxed{\frac{1}{4}},$$

$$v_1^{\sigma^1}(I_K, \text{bet}) = \left(\frac{1}{6} + \frac{1}{6}\right) \left[\frac{1}{2}(1) + \frac{1}{2}(2)\right] = \boxed{\frac{1}{2}},$$

$$v_1^{\sigma^1}(I_K) = \frac{1}{2} \left(\frac{1}{4}\right) + \frac{1}{2} \left(\frac{1}{2}\right) = \boxed{\frac{3}{8}}.$$

Counterfactual value is not a conditional payoff

The per-node conditional expectation is $\frac{9}{8}$; multiplying by the total counterfactual reach $\frac{1}{3}$ gives $\frac{3}{8}$.

L13 connection: The outer sum supplies the chance-and-opponent weighting in $v_i^{\sigma}(I)$.

The first regret-matching update at P1's J

L13's instantaneous regret compares a forced action with the current mixture:

$$r_i^t(I, a) = v_i^{\sigma^t}(I, a) - v_i^{\sigma^t}(I).$$

$$r_1^1(I_J, \text{bet}) = -\frac{1}{6} - \left(-\frac{7}{24}\right) = \boxed{\frac{1}{8}},$$

$$r_1^1(I_J, \text{check}) = -\frac{5}{12} - \left(-\frac{7}{24}\right) = \boxed{-\frac{1}{8}}.$$

Start with $R_i^0(I, a) = 0$, so $R_1^1(I_J, a) = r_1^1(I_J, a)$. For the *next* iteration, normalize the positive cumulative regrets:

$$\sigma_1^2(\text{bet} \mid I_J) = \frac{1/8}{0 + 1/8} = \boxed{1}, \quad \sigma_1^2(\text{check} \mid I_J) = \boxed{0}.$$

If all cumulative regrets are nonpositive, use $1/2, 1/2$. **Keep negative regrets in memory;** only the matching step takes positive parts.

L13 connection: Instantaneous regret r_i^t , cumulative regret R_i^t , and regret matching.

All three opening information sets update together

All values below use the *same frozen* uniform profile σ^1 .

P1 set	$v(\text{check})$	$v(\text{bet})$	$v(I)$	$R^1(\text{check})$	$R^1(\text{bet})$	$\sigma_1^2(\text{bet} \mid I)$
I_J	$-\frac{5}{12}$	$-\frac{1}{6}$	$-\frac{7}{24}$	$-\frac{1}{8}$	$\frac{1}{8}$	1
I_Q	$-\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{24}$	$-\frac{1}{8}$	$\frac{1}{8}$	1
I_K	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{8}$	$-\frac{1}{8}$	$\frac{1}{8}$	1

Betting all three cards is a first update, not an equilibrium. The opponents' next strategies change the values in the next round.

The iteration convention matters

Compute every player's values under σ^t , update all regrets, then form σ^{t+1} . No strategy changes during the traversal.

We enumerate all six deals on every iteration and keep every action available.

L13 connection: The simultaneous-update convention in L13's vanilla CFR pseudocode.

The decomposition theorem, with actual totals

For a fixed complete strategy deviation, L13's game-level regret is bounded by the sum of the largest positive local regrets:

$$R_i^T = \max_{\sigma'_i} \sum_{t=1}^T [u_i(\sigma'_i, \sigma_{-i}^t) - u_i(\sigma^t)] \leq \sum_{I \in \mathcal{I}_i} \max_a R_i^{T,+}(I, a).$$

Preview of the next three rounds: enumerate all 64 pure plans per player.

At $T = 3$	Global external regret	Sum of local positive maxima
Player 1	$\frac{13}{12}$	$\frac{13}{12}$
Player 2	$\frac{13}{24}$	$\frac{13}{24}$

Here the bound holds with equality. Local regret matching is separate at each I , but the values couple those decisions through the game tree.

L13 connection: CFR decomposition: bound whole-strategy regret using local counterfactual regrets.

Iteration $t = 1$: all 12 information sets

$a_0/a_1 = x/b$ at $\emptyset, x$; $a_0/a_1 = f/c$ at xb, b . $p = \sigma_i^t(a_1 | I)$, $p' = \sigma_i^{t+1}(a_1 | I)$; the other action has probability $1 - p$.

Abbreviate $v(a) = v_i^{\sigma^t}(I, a)$, $v = v_i^{\sigma^t}(I)$, and $R(a) = R_i^t(I, a)$.

Player 1

I	p	$v(a_0)$	$v(a_1)$	v	$R(a_0)$	$R(a_1)$	p'
$I_{J,\emptyset}^1$	1/2	-5/12	-1/6	-7/24	-1/8	1/8	1
$I_{Q,\emptyset}^1$	1/2	-1/12	1/6	1/24	-1/8	1/8	1
$I_{K,\emptyset}^1$	1/2	1/4	1/2	3/8	-1/8	1/8	1
$I_{J,xb}^1$	1/2	-1/6	-1/3	-1/4	1/12	-1/12	0
$I_{Q,xb}^1$	1/2	-1/6	0	-1/12	-1/12	1/12	1
$I_{K,xb}^1$	1/2	-1/6	1/3	1/12	-1/4	1/4	1

Player 2

I	p	$v(a_0)$	$v(a_1)$	v	$R(a_0)$	$R(a_1)$	p'
$I_{J,x}^2$	1/2	-1/6	-1/12	-1/8	-1/24	1/24	1
$I_{Q,x}^2$	1/2	0	1/12	1/24	-1/24	1/24	1
$I_{K,x}^2$	1/2	1/6	1/4	5/24	-1/24	1/24	1
$I_{J,b}^2$	1/2	-1/6	-1/3	-1/4	1/12	-1/12	0
$I_{Q,b}^2$	1/2	-1/6	0	-1/12	-1/12	1/12	1
$I_{K,b}^2$	1/2	-1/6	1/3	1/12	-1/4	1/4	1

$R^1 = r^1$. Each action value includes chance; rows use the acting player's utility.

L13 connection: Values, cumulative regrets, and the next strategy follow the simultaneous-update pseudocode.

Iteration $t = 2$: all 12 information sets

$a_0/a_1 = x/b$ at $\emptyset, x$; $a_0/a_1 = f/c$ at xb, b . $p = \sigma_i^t(a_1 | I)$, $p' = \sigma_i^{t+1}(a_1 | I)$; the other action has probability $1 - p$.

Abbreviate $v(a) = v_i^{\sigma^t}(I, a)$, $v = v_i^{\sigma^t}(I)$, and $R(a) = R_i^t(I, a)$.

Player 1

I	p	$v(a_0)$	$v(a_1)$	v	$R(a_0)$	$R(a_1)$	p'
$I_{J,\emptyset}^1$	1	-1/3	-2/3	-2/3	5/24	1/8	3/8
$I_{Q,\emptyset}^1$	1	0	-1/6	-1/6	1/24	1/8	3/4
$I_{K,\emptyset}^1$	1	2/3	1/2	1/2	1/24	1/8	3/4
$I_{J,xb}^1$	0	-1/3	-2/3	-1/3	1/12	-5/12	0
$I_{Q,xb}^1$	1	-1/3	0	0	-5/12	1/12	1
$I_{K,xb}^1$	1	-1/3	2/3	2/3	-5/4	1/4	1

Player 2

I	p	$v(a_0)$	$v(a_1)$	v	$R(a_0)$	$R(a_1)$	p'
$I_{J,x}^2$	1	0	0	0	-1/24	1/24	1
$I_{Q,x}^2$	1	0	0	0	-1/24	1/24	1
$I_{K,x}^2$	1	0	0	0	-1/24	1/24	1
$I_{J,b}^2$	0	-1/3	-2/3	-1/3	1/12	-5/12	0
$I_{Q,b}^2$	1	-1/3	0	0	-5/12	1/12	1
$I_{K,b}^2$	1	-1/3	2/3	2/3	-5/4	1/4	1

P1 never checks in σ^2 , so P2's x rows have zero CF value. P1's xb rows still update!

L13 connection: Values, cumulative regrets, and the next strategy follow the simultaneous-update pseudocode.

Iteration $t = 3$: all 12 information sets

$a_0/a_1 = x/b$ at $\emptyset, x$; $a_0/a_1 = f/c$ at xb, b . $p = \sigma_i^t(a_1 \mid I)$, $p' = \sigma_i^{t+1}(a_1 \mid I)$; the other action has probability $1 - p$.

Abbreviate $v(a) = v_i^{\sigma^t}(I, a)$, $v = v_i^{\sigma^t}(I)$, and $R(a) = R_i^t(I, a)$.

Player 1

I	p	$v(a_0)$	$v(a_1)$	v	$R(a_0)$	$R(a_1)$	p'
$I_{J,\emptyset}^1$	3/8	-1/3	-2/3	-11/24	1/3	-1/12	0
$I_{Q,\emptyset}^1$	3/4	0	-1/6	-1/8	1/6	1/12	1/3
$I_{K,\emptyset}^1$	3/4	2/3	1/2	13/24	1/6	1/12	1/3
$I_{J,xb}^1$	0	-1/3	-2/3	-1/3	1/12	-3/4	0
$I_{Q,xb}^1$	1	-1/3	0	0	-3/4	1/12	1
$I_{K,xb}^1$	1	-1/3	2/3	2/3	-9/4	1/4	1

Player 2

I	p	$v(a_0)$	$v(a_1)$	v	$R(a_0)$	$R(a_1)$	p'
$I_{J,x}^2$	1	-1/12	-1/6	-1/6	1/24	1/24	1/2
$I_{Q,x}^2$	1	1/16	1/48	1/48	0	1/24	1
$I_{K,x}^2$	1	7/48	3/16	3/16	-1/12	1/24	1
$I_{J,b}^2$	0	-1/4	-1/2	-1/4	1/12	-2/3	0
$I_{Q,b}^2$	1	-3/16	-1/8	-1/8	-23/48	1/12	1
$I_{K,b}^2$	1	-3/16	3/8	3/8	-29/16	1/4	1

For I_J , $R^3 = (1/3, -1/12)$, so the next strategy checks with probability 1.

L13 connection: Values, cumulative regrets, and the next strategy follow the simultaneous-update pseudocode.

Average the strategy with the player's own reach

L13 accumulates $S_i(l, a) \leftarrow S_i(l, a) + \pi_i^{\sigma^t}(l) \sigma_i^t(a | l)$:

$$\bar{\sigma}_i^T(a | l) = \frac{\sum_{t=1}^T \pi_i^{\sigma^t}(l) \sigma_i^t(a | l)}{\sum_{t=1}^T \pi_i^{\sigma^t}(l)}.$$

At the opening l_J

Own reach is 1 on every round:

$$\bar{\sigma}_1^3(\text{bet} | l_J) = \frac{1/2 + 1 + 3/8}{3} = \boxed{\frac{5}{8}}.$$

At $l_{Q,xb}^1$

Own reach is P1's earlier check probability:

t	1	2	3
$\pi_1^{\sigma^t}(l)$	$\frac{1}{2}$	0	$\frac{1}{4}$
$\sigma_1^t(c l)$	$\frac{1}{2}$	1	1

$$\bar{\sigma}_1^3(c | l_{Q,xb}^1) = \frac{(1/2)(1/2) + (0)(1) + (1/4)(1)}{1/2 + 0 + 1/4} = \boxed{\frac{2}{3}}.$$

Averaging uses own reach. A plain average would incorrectly give 5/6 here.

L13 connection: Regret updates use opponent/chance reach; average-strategy updates use own reach.

After three rounds: improvement, but far from equilibrium

Each cell shows $\bar{\sigma}^3(a \mid I) \mid \sigma^*(a \mid I)$ for the named action.

Information-set family	Action	J	Q	K
$I_{d,\emptyset}^1$	bet	5/8 1/3	3/4 0	3/4 1
$I_{d,xb}^1$	call	2/9 0	2/3 2/3	2/3 1
$I_{d,x}^2$	bet	5/6 1/3	5/6 0	5/6 1
$I_{d,b}^2$	call	1/6 0	5/6 1/3	5/6 1

What improved? The sum of best-response gains (NashConv) fell: $\frac{11}{12} \longrightarrow \frac{5}{8} \longrightarrow \frac{13}{24}$.

J bluff: $1/2 \rightarrow 5/8$. Q call: $1/2 \rightarrow 5/6$.

Both moved away from 1/3. The game-level gap can improve even when individual coordinates do not; P1 may approach a different equilibrium.

L13 connection: The average profile is the candidate approximate NE; assess it using best responses.

Current play changes abruptly; averaging smooths it

For P1's opening J , own reach is 1, so the weighted average is an ordinary mean.

t	$\sigma_1^t(\text{bet} \mid I_J)$	$\bar{\sigma}_1^t(\text{bet} \mid I_J)$	$\sigma_2^t(\text{call} \mid I_{Q,b}^2)$	$\bar{\sigma}_2^t(\text{call} \mid I_{Q,b}^2)$
1	1/2	1/2	1/2	1/2
2	1	3/4	1	3/4
3	3/8	5/8	1	5/6
4	0	15/32	1	7/8
5	0	3/8	1	9/10

The current J strategy jumps; its average moves more gradually. Smoothing need not improve every coordinate immediately. **Guarantee:** In two-player zero-sum games with perfect recall, CFR's average approaches the NE set. There is no general last-iterate guarantee.

L13 connection: L13's average-strategy guarantee concerns exploitability, not monotone coordinate motion.

A longer run: measure convergence by deviation gains

Define the sum of unilateral best-response gains:

$$\text{NashConv}(\sigma) = \underbrace{\max_{\sigma'_1} u_1(\sigma'_1, \sigma_2) - u_1(\sigma)}_{\text{P1's gain}} + \underbrace{\max_{\sigma'_2} u_2(\sigma_1, \sigma'_2) - u_2(\sigma)}_{\text{P2's gain}}.$$

T	P1 average J bet	P1 average K bet	P2 average Q call	$\text{NashConv}(\bar{\sigma}^T)$
10	0.187500	0.681542	0.892002	0.192417
100	0.194223	0.550969	0.436274	0.051349
1000	0.201009	0.598473	0.368790	0.014538
10000	0.207755	0.612593	0.339285	0.004636

P2's Q call approaches 1/3. P1's J bet approaches an equilibrium-family value α , with K bet near 3α ; it need not approach 1/3.

At $T = 10000$, $u_1(\bar{\sigma}^T) \approx -0.055546$; the exact game value is $-1/18$.

L13 connection: The average-regret bound gives convergence to the Nash equilibrium set.

From local arithmetic to a game-level equilibrium

We traced CFR through Kuhn poker:

reach $\longrightarrow$ counterfactual values $\longrightarrow$ regret
 $\longrightarrow$ regret matching $\longrightarrow$ iteration $\longrightarrow$ average-strategy convergence.

- **12 information sets, two actions each:** 12 local regret-matching problems and 24 cumulative regret entries.
- Local updates are simple; their counterfactual values depend on the full tree.
- The complete reference equilibrium has value $u_1 = -1/18$ and zero best-response gains. The computed average approaches the equilibrium set.
- Three rounds expose the mechanism. A longer run demonstrates the decrease in best-response gains.

L13 connection: The decomposition theorem connects the local updates to global regret and NE.

Sources and reproducibility

- Alan Kuhnle, *Lecture 13: Counterfactual Regret Minimization*. This deck preserves its notation and simultaneous-update convention.
- Martin Zinkevich, Michael Johanson, Michael Bowling, Carmelo Piccione (2007, NIPS), *Regret Minimization in Games with Incomplete Information*.
- Bret Hoehn, Finnegan Southey, Robert C. Holte, Valeriy Bulitko (2005, AAI), *Effective Short-Term Opponent Exploitation in Simplified Poker*: Kuhn rules and equilibrium family.

Independent checks: exact rational tree traversal versus terminal enumeration; NumPy reproduction; 64 pure deviations per player; a separate normal-form LP. All numerical trace entries in the TeX source carry verification keys.

```
python3 /tmp/kuhn_cfr_verify.py
```

First five iterations are exact; longer-run decimals are rounded to six places. The PDF compiles from this TeX file and 631-preamble.tex alone.

L13 connection: The verification follows L13's definitions directly, including off-path information sets.