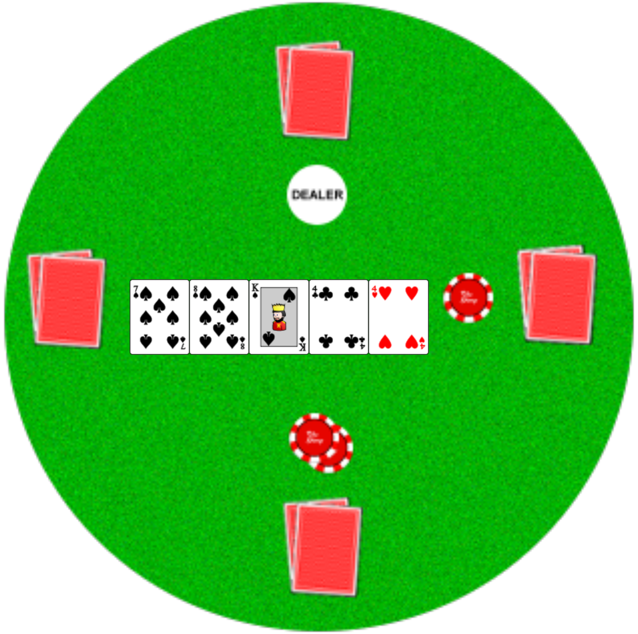
Heads-up limit hold'em poker is solved

Background Presentation

Texas Hold’em

Rounds:

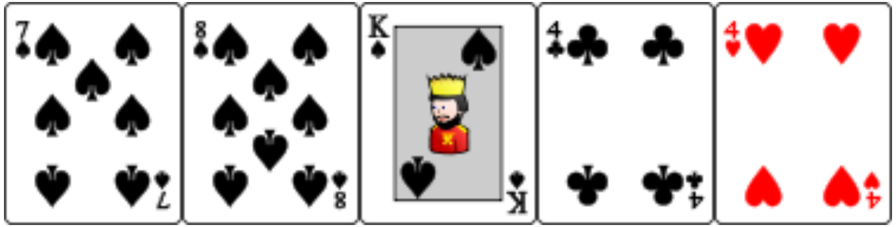
- hole cards
- betting
- flop
- betting
- turn
- betting
- river
- betting
- showdown



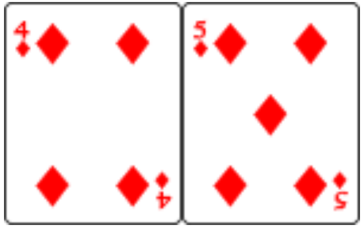
Betting

- check
- fold
- call
- raise (amount)

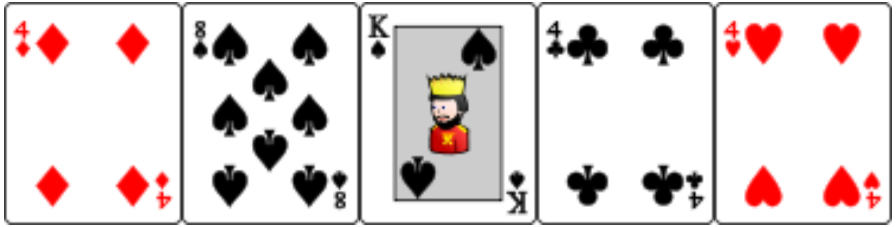
Showdown



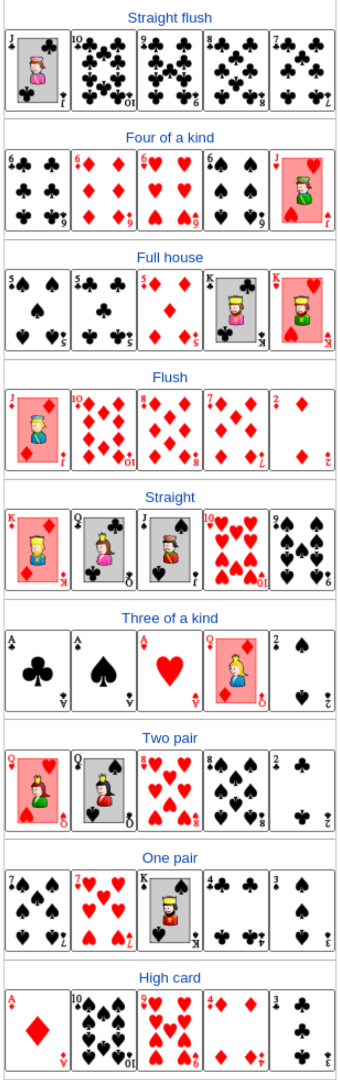
common



hole



best 5-card hand



HUHLE

Two major simplifications:

- Heads-up: 2 players
  - zero-sum
- Limit: fixed betting increment
  - much smaller strategy space

But people actually play this game!

Value of a Game

According to the minimax theorem, in 2-player zero-sum games:

- Maximizing your own payoff is equivalent to minimizing your opponent's.
- Maximizing your worst-case payoff results in a NE strategy.
- All Nash equilibria have the same payoffs.

The NE payoff to player 1 is called the value of the game.

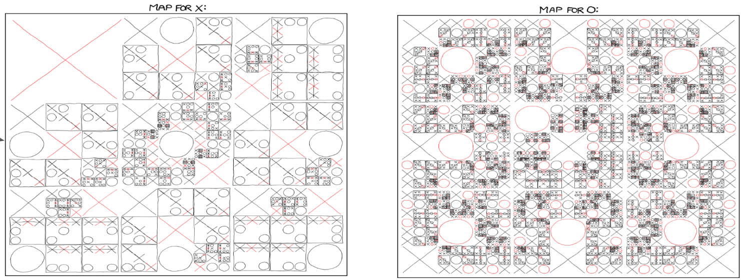
The value of rock-paper-scissors is 0.

Strength of a Game Solution

- ultra-weakly solved
  - The value of the game is known (but strategies to achieve it are not).
- weakly solved
  - A Nash equilibrium strategy is known (but off-path optimal play is not).
- strongly solved
  - Optimal play is known from anywhere in the game tree.

Hex

Checkers



Tic-Tac-Toe

Unsolved Games

- computers better than any human
  - chess

- computers are worse than the best professionals
  - go

- computers are worse than many amateurs
  - n-player no-limit hold’em

The State of HUHLE

Polaris lost to human professionals in 2007, but won in 2008.

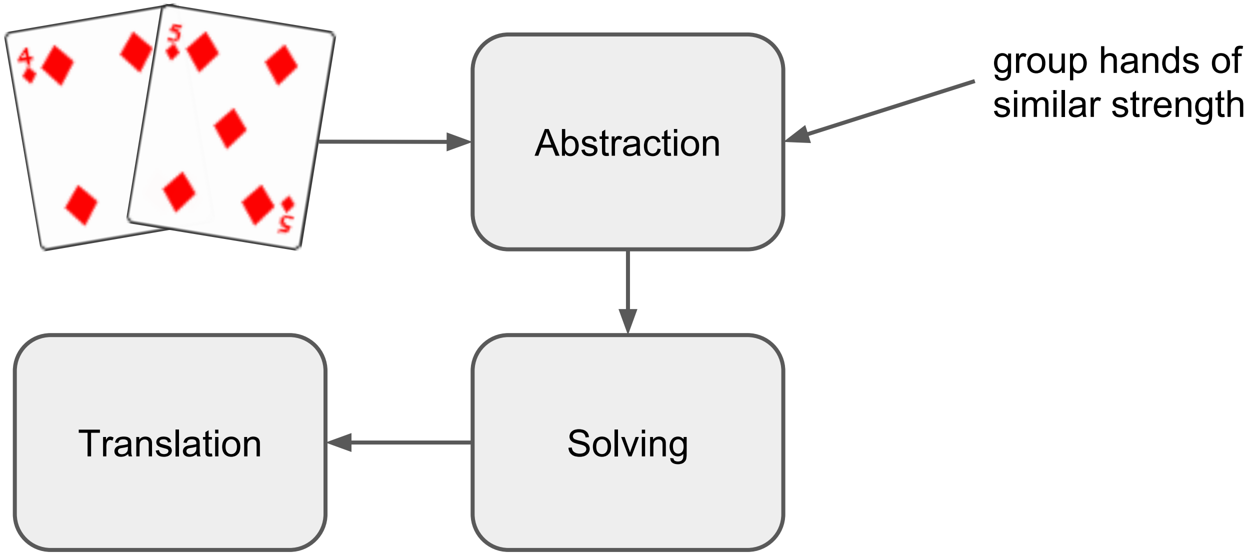
Human-computer tournaments are hard to run because a huge number of games are required for statistical significance.

Polaris is now known to be exploitable for roughly ¼ big bet per game.

Size of the Game Tree

HUHLE has 1.38×10^{13} information sets.

The largest previously solved abstraction has 3.8×10^{10} information sets.



Big Developments in Solving HUHLE

- *Regret Minimization in Games with Incomplete Information. Zinkevich, Johanson, Bowling, and Piccione. NIPS 2009.*
  - Self-play algorithm that converges to NE
  - Can run on larger abstractions than previous algorithms

- *Accelerating Best Response Calculation in Large Extensive Games. Johanson, Waugh, Bowling, and Zinkevich. IJCAI 2011.*
  - Allows exploitability (regret) calculations

Counterfactual Regret Minimization (CFR)

Key Ideas:

- Iteratively improve strategies through self-play.
  - Reduce regret on each iteration.
- Split up regret into independent additive terms.
  - Counterfactual regret value for each information set.
  - Sum of CFR values bounds total regret.
- CFR is (roughly) the expected gain from switching one action.
- Choose actions to minimize CFR at each information set.

Accelerated Best-Response Calculation

Key Ideas:

- Efficiently re-use information from the public game tree & opponent strategy.
  - Requires re-ordering computation to evaluate game tree nodes with the same public information together.
- Exploit the ranking of hands in expected value computation.
  - Don’t need to compute EV for all opponent information sets; two hands that you beat have the same EV.
- Suit isomorphisms
  - Swap all hearts for clubs and the outcome is the same.
- Parallel Computation
  - Split up independent subtrees.