A Memory-Based Approach to Two-Player Texas Hold'em

Jonathan Rubin & Ian Watson
University of Auckland, NZ

http://www.cs.auckland.ac.nz/research/gameai

Introduction

- Memory-Based Approach
  - Simple approach
- Produce poker strategy
- Agent
  - Sartre

Overview

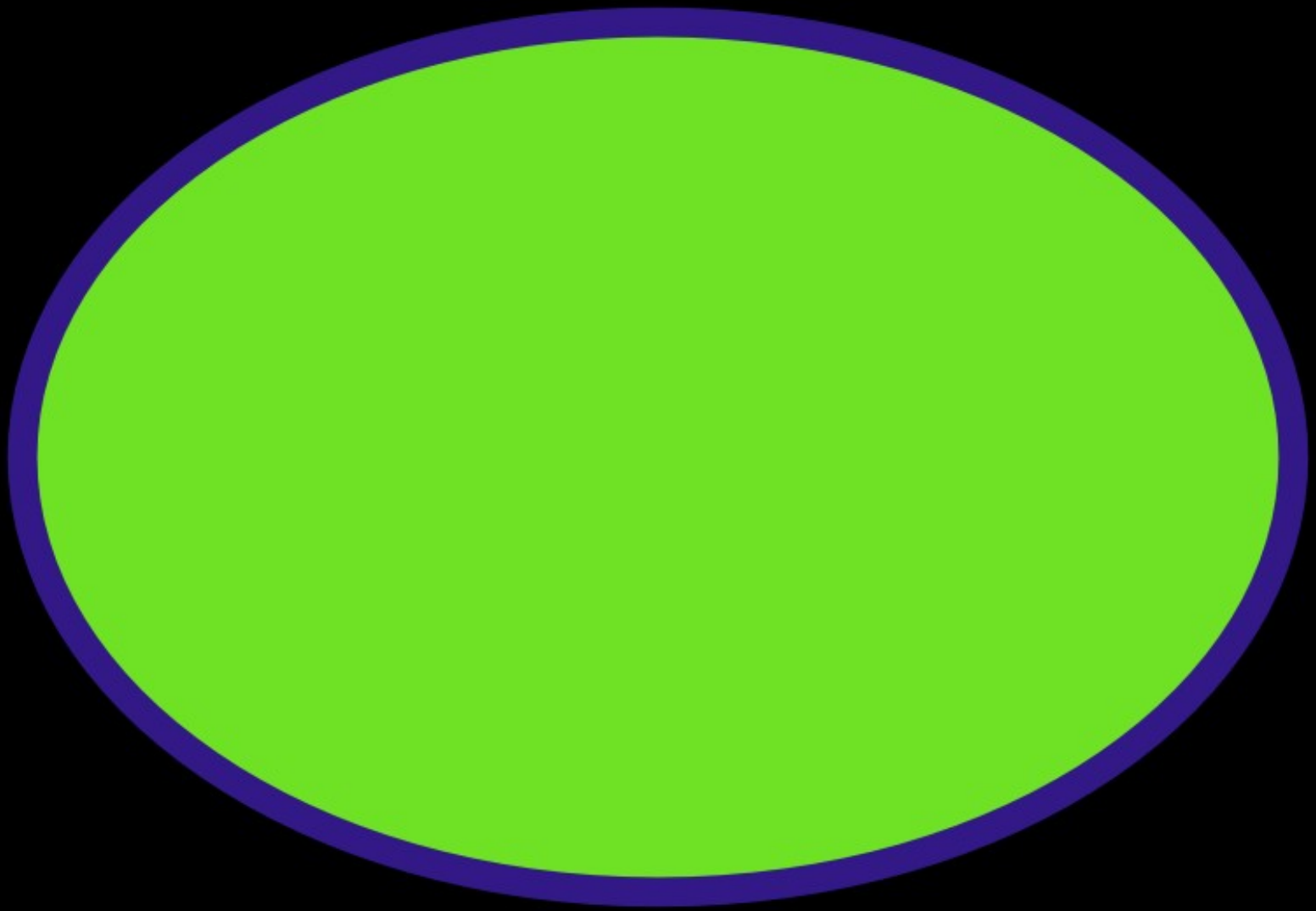
- Texas Hold'em
- Types of Strategies
- Related Approaches
- Memory-Based Approach
- Experimental Results
- Conclusions
- Future Work

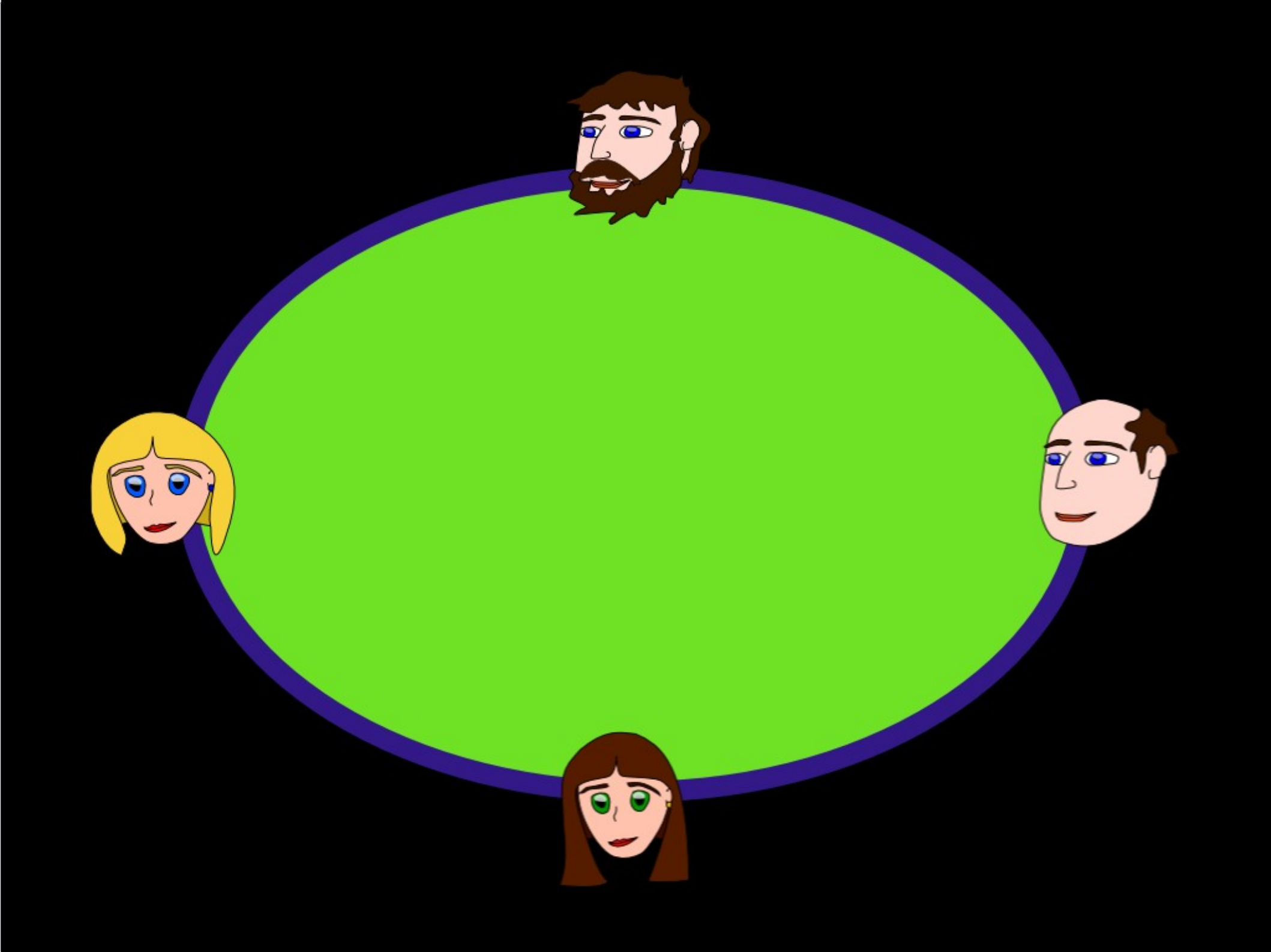
Introduction

The Poker Domain

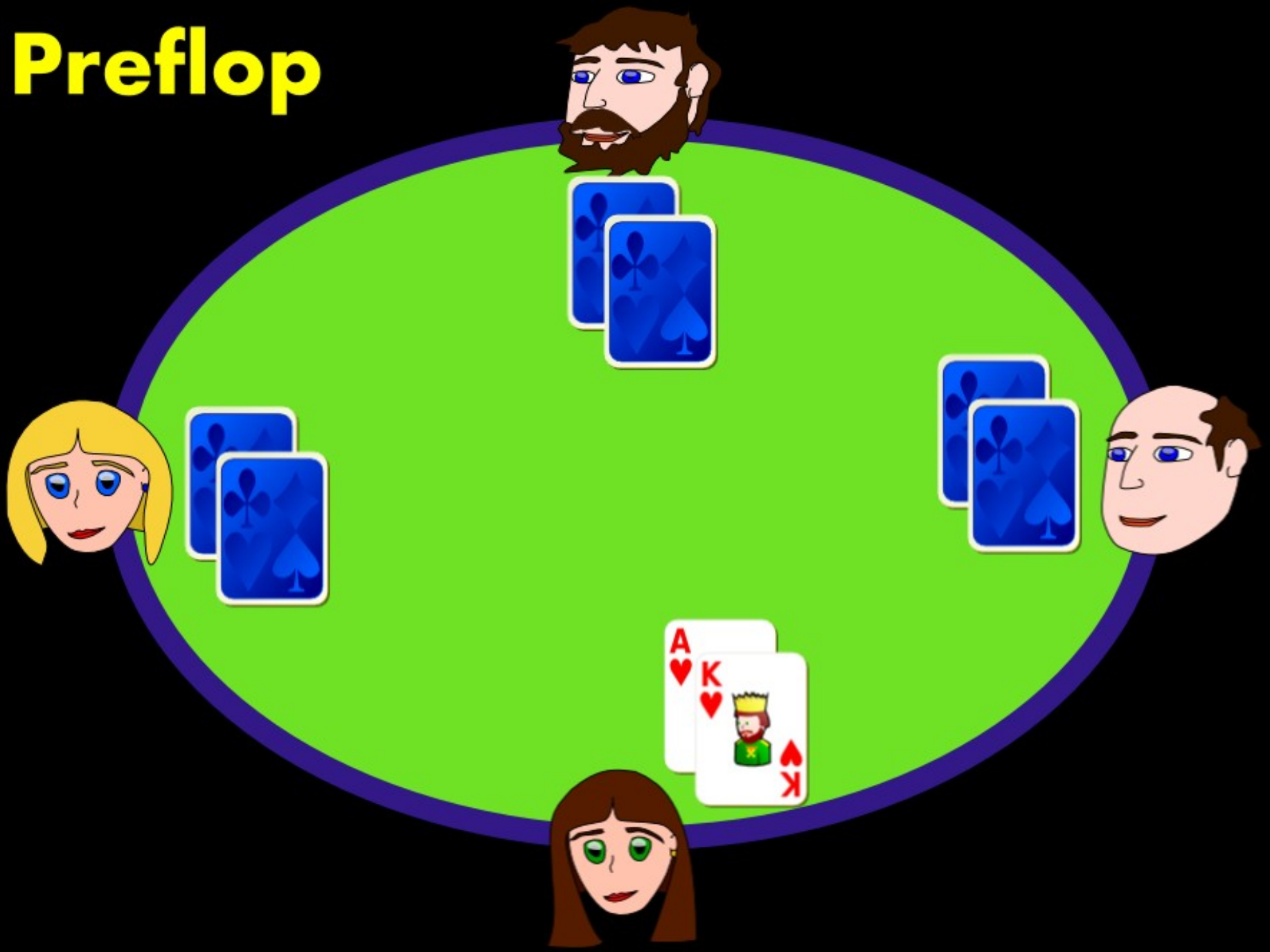
- Imperfect Information
- Chance events
- Rules and boundaries
- Performance evaluation
- Increasingly popular
  - AAAI Computer Poker Competition

The Rules of Texas Hold'em





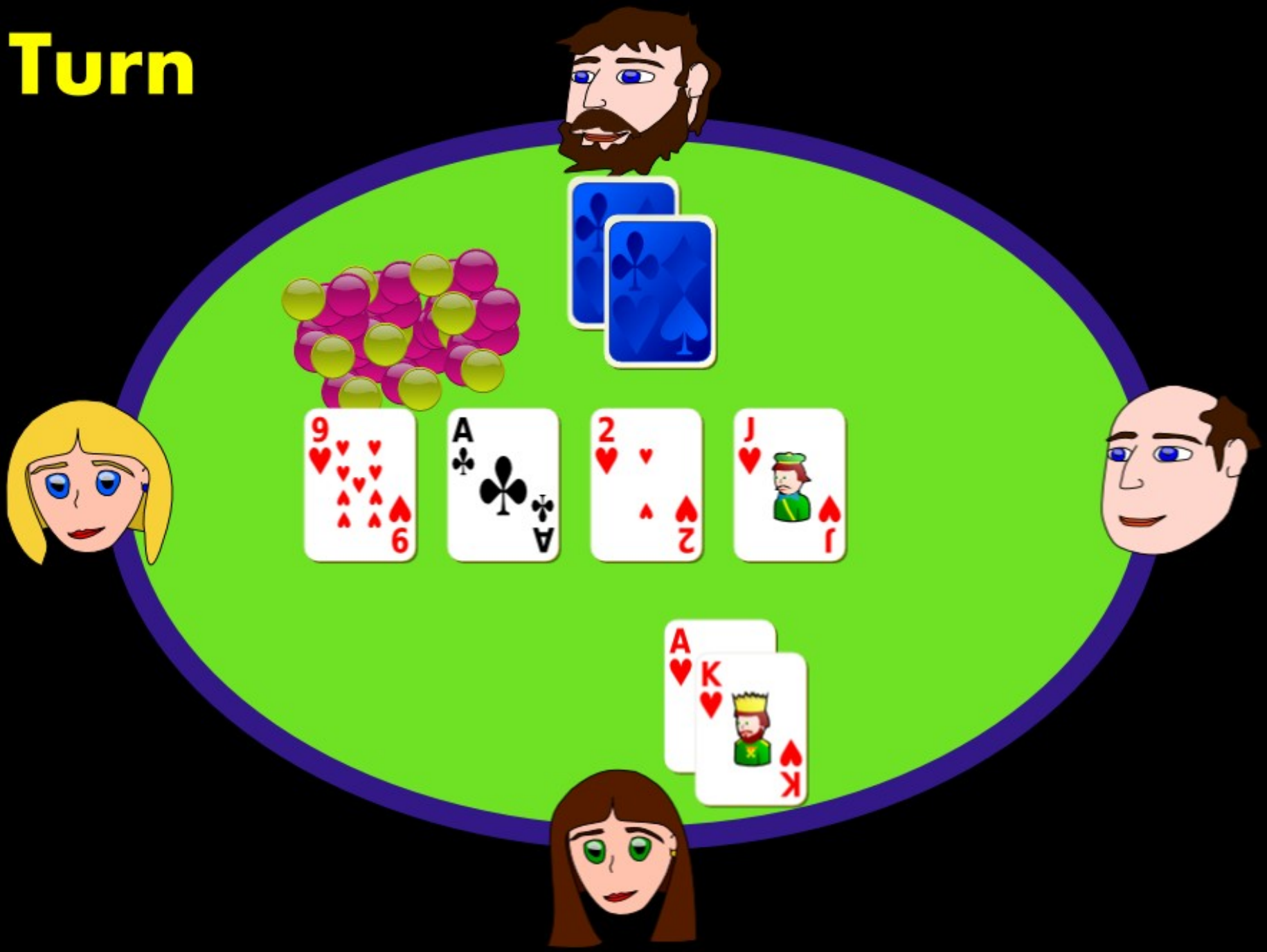
Preflop



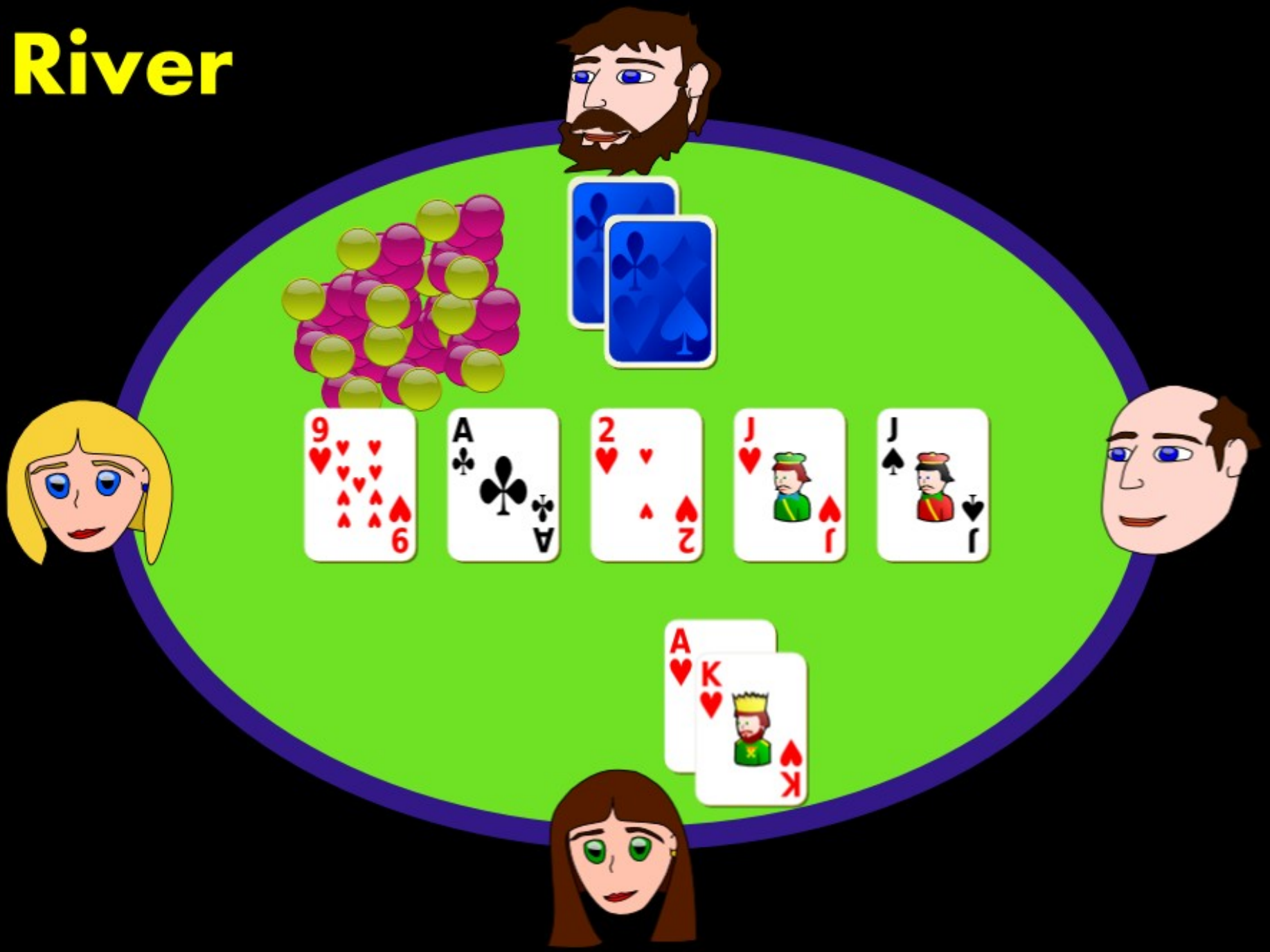
Preflop

Flop

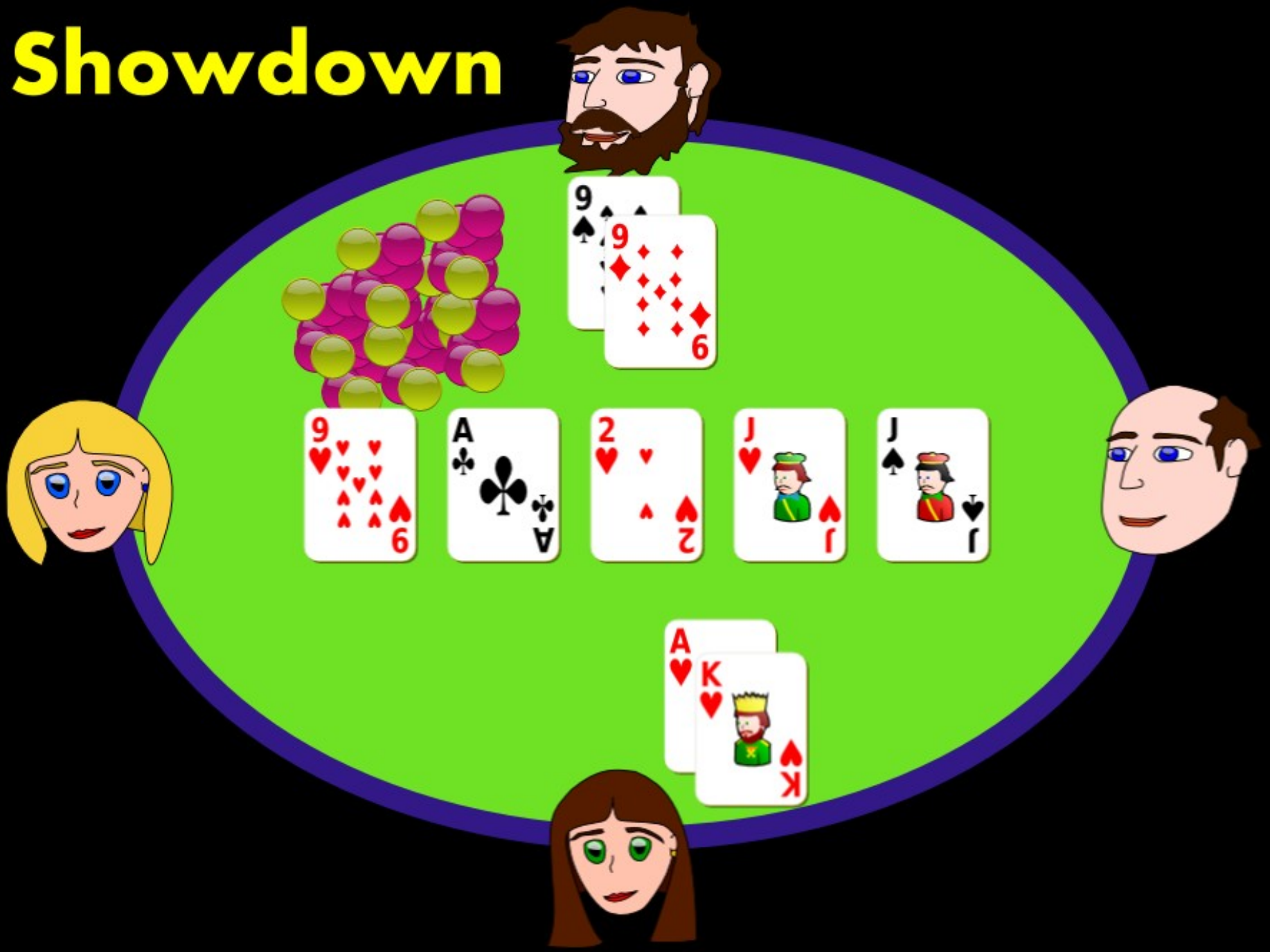
Turn



River



Showdown



Texas Hold'em

- Current Focus
  - Heads up (2 players)
  - Limit betting
    - $2/$4 Hold'em

Poker Strategies

A Poker Strategy

- At every decision point a probability triple is required that indicates the proportion of the time a player should either fold, call or raise

$$(f,c,r) \rightarrow (0,\ 0.5,\ 0.5)$$

Types of Strategies

- Nash Equilibria
  - Robust strategies that attempt not to lose to any type of opponent
- Exploitive Strategies
  - Attempts to react to an opponent's play in a way that allows maximum exploitability of that opponent
  - Requires opponent modeling

Rock-Paper-Scissors Example

- Nash equilibrium
  - *(R,P,S) → (1/3, 1/3, 1/3)*
  - The Nash player will never lose against any player in the long run
- Along comes Jimmy who only ever plays Paper

Rock-Paper-Scissors Example

- The Nash player will continue to play
  - *(1/3, 1/3, 1/3)*
  - Lose 33%, Win 33%, Draw 33%
  - The Nash player will still only draw against Jimmy

Rock-Paper-Scissors Example

- However, because we know Jimmy's strategy an exploitive player would be better off using the strategy
  - *(0, 0, 1.0)*
  - i.e. a best response that maximally exploits Jimmy at every decision point
- Now, against Jimmy the exploitive player will win
  - Consequence is that the exploitive player plays off the equilibrium, and is hence subject to potential exploitation itself

Approaches to creating poker agents

e-Nash Equilibrium

- Linear Programming
  - Constructs matrices that act as constraints within an optimization problem
- Iterative approaches
  - Basic idea: Two players begin with arbitrary strategies, play many repetitions of a game and modify their strategies in a way that improves their strategy against their opponent.
  - As the number of iterations increases the strategies approach a Nash equilibrium
  - e.g. *Fictitious Play*, *Counterfactual Regret Minimization*

e-Nash Equilibrium

- A Nash equilibrium can easily be computed for Rock-Paper-Scissors
- However, the poker game tree is much to large to find exact Nash equilibria
  - Abstractions required
- Can only approximate Nash-equilibria
  - *e*-Nash Equilibria
  - *e* specifies a lower bound on how exploitable the equilibrium strategy is

Exploitive Strategies

- Miximax search
  - Similar to minimax in perfect information games
  - Maintains an opponent model used during game tree search to inform expected value calculations of taking certain betting actions
- Restricted Nash Response (RNR) & Data Biased Response (DBR)
  - Somewhere between an *e*-Nash equilibrium and a best response to an opponent's static strategy

Our Approach

Goal

- Investigate whether hand histories from strong poker players can be reused within a Case-Based Reasoning framework to achieve a similar performance?

Case-Based Reasoning (CBR)

- Solutions of past problems are reused or adapted to handle solutions for novel problems
- Lazy Learning approach
- Stores a memory of cases along with their solutions and outcomes
- When a new problem is encountered similar cases are retrieved from the case-base and their solutions are reused to solve the problem

A Memory-Based Approach

- Casper (CASe-based Poker playER)
  - Past poker agent for 10-player Texas Hold'em
- Sartre (Similarity Assessment Reasoning for Texas hold'em via Recall of Experience)
  - Our latest agent
  - Specialised for heads-up limit hold'em

Sartre

- Overview
  - Cases are attribute-value pairs
  - Separate case-bases are used for each different round (preflop, flop, turn, river)
  - When a decision is required a case is created to describe the current state of the game and the appropriate case-base is searched to find similar cases
  - The solution of the similar cases are reused for the current situation

Sartre

- Case Representation

Attribute	Type	Example
Hand Type	Class	*Missed, Pair, Two-Pair, Set, Flush,Flush-Draw, Straight-Draw, ...*
Betting Sequence	String	*rc-c, crrc-crrc-cc-r, ...*
Board Texture	Class	*No-Salient, Flush-Possible, Straight-Possible, Flush-Highly-Possible, ...*
Solution	Char	*f, c, r*
Outcome	Numerical	*+14, -1, -5, +20, ...*

Sartre

- Case Retrieval
  - Current version of Sartre uses All-or-Nothing local similarity, i.e. either attribute values are entirely similar or dissimilar
    - Baseline for future improvements
  - Number of retrieved cases varies from 0 to 1000s
  - If 0 cases retrieved Sartre adopts a default strategy
    - Allways-Call

Sartre

- Solution Reuse
  - Many cases retrieved – which betting action to make?
  - 3 solution reuse policies
    - 1) Reuse the majority decision
    - 2) Probabilistically select actions*
    - 3) Reuse solution which achieved the greatest outcome

Sartre

- Training Data
  - Trained on data from the best agent equilibrium agent from the 2008 Computer Poker Competition
    - Hyperborean-Eqm

Round	# of Cases
Preflop	201,335
Flop	300,577
Turn	281,529
River	216,597

Experimental Results

Experiments

- Where possible used: Duplicate Matches
  - *N* hands in forward + backwards direction
  - Set of hands played
  - Set of hands replayed, but agents receive the cards that their opponent previously received
  - Reduces variance
- Small bets per hand (sb/h)

Experiments

- Sartre Vs. FellOmen2
  - Sartre “expert” trained by Hyperborean-Eqm
  - Hyperborean-Eqm Vs. FellOmen2 results known
  - Compare Sartre Vs. FellOmen2 to Hyperborean Vs. FellOmen2
  - FellOmen2 2nd equal in 2008 equilibrium CPC
  - Publicly available
- 6 rounds of $N = 3000$ duplicate hands

Experiments

- Sartre Vs. FellOmen2

Round	Sartre (sb/h)	Hyperborean (sb/h)*
1	-0.025	+0.014
2	-0.041	+0.023
3	-0.094	+0.029
4	-0.055	+0.030
5	-0.066	+0.033
6	-0.070	+0.016
Average	**-0.0585 +/- 0.01 sb/h**	**+0.0241 +/- 0.003 sb/h**

* Note: $N = 5000$

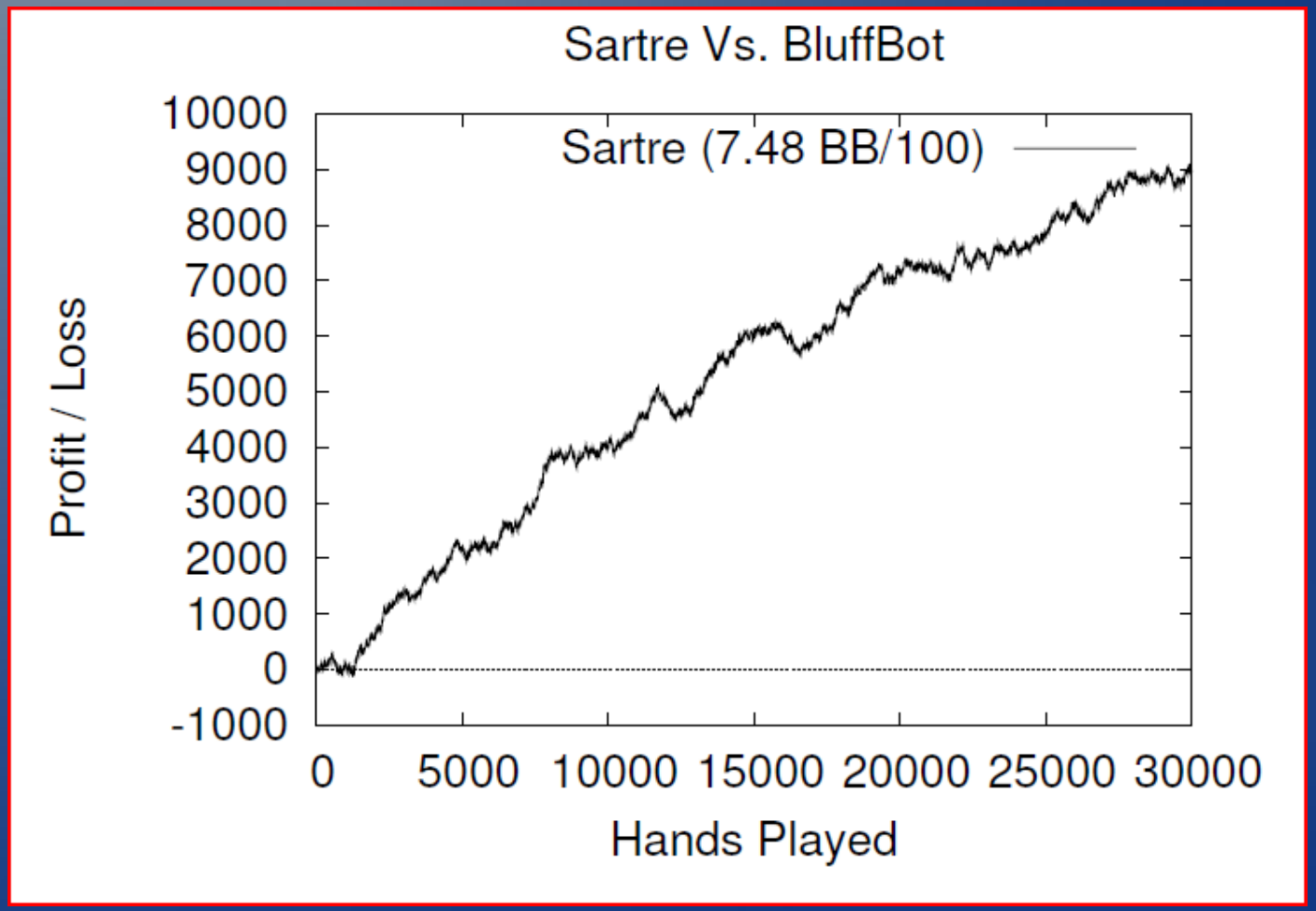
- Independent samples t-test gives $p < 0.00001$

Experiments

- Sartre Vs. BluffBot
  - Further evaluation
  - 2nd place in 2006 Computer Poker Competition
  - Publicly available
  - Duplicate match structure not available
  - Straight 30,000 hands

Experiments

- Sartre Vs. BluffBot
  - +0.150 sb/h



Experiments

- 2009 IJCAI Computer Poker Competition
  - Participated in limit hold'em competition
  - Same system, but with majority-decision reuse policy
    - Chosen because of results of self-play experiments
  - 13 competitors
  - 2 divisions
    - Bankroll
    - Equilibrium

Experiments

- 2009 IJCAI Computer Poker Competition
  - Limit bankroll division

Place	Agent	sb/h
1	MANZANA	0.186
2	Hyperborean-BR	0.116
3	GGValuta	0.110
4	Hyperborean-Eqm	0.116
5	Rockhopper	0.103
6	**Sartre**	**0.097**
7	Slumbot	0.096
8	GS5	0.082
9	AoBot	-0.002
10	dcurbHU	-0.07
11	LIDIA	-0.094
12	GS5Dynamic	-0.201

Experiments

- 2009 IJCAI Computer Poker Competition
  - Limit equilibrium division

Place	Agent
1	GGValuta
2	Hyperborean-Eqm
3	MANZANA
4	Rockhopper
5	Hyperborean-BR
6	Slumbot
7	**Sartre**
8	GS5
9	AoBot
10	GS5Dynamic
11	LIDIA
12	dcurbHU
13	Tommybot

Conclusions

Conclusions

- Presented a straight-forward, memory based approach for 2-player limit Texas Hold'em
- Initial results show a disparity between our memory-based system trained via “expert” and actual “expert” player results
  - With further improvements we believe we can limit this gap
- Memory-based approach still able to achieve strategies of reasonable quality
  - Consistent profit against BluffBot
  - 6th place finish in 2009 Computer Poker Competition

Future Work

Future Work

- Improved similarity measures & generalization
  - Compare against Sartre-Baseline
- Investigate Case Representation
- No limit betting

Thank you!

To challenge Sartre go to:

www.cs.auckland.ac.nz/poker