Robust Strategies and Counter-Strategies
Building a Champion Level Computer Poker Player

Mike Johanson

November 20, 2012

University of Alberta
Computer Poker Research Group
How can we create a poker program for competing against expert players?
How can we create a poker program for competing against expert players?

- Three new techniques for finding game theoretic strategies
How can we create a poker program for competing against expert players?

- Three new techniques for finding game theoretic strategies
- Useful for poker, applicable to other domains

Show the value of these approaches through competitions against expert humans and computers.
How can we create a poker program for competing against expert players?

- Three new techniques for finding game theoretic strategies
- Useful for poker, applicable to other domains
- Show the value of these approaches through competitions against expert humans and computers
1 Introduction

2 Playing to Not Lose: Counterfactual Regret Minimization

3 Playing to Win: Frequentist Best Response

4 Playing to Win, Carefully: Restricted Nash Response

5 Competition Results

6 Conclusion
The CPRG’s goal: Create poker programs to beat the world’s best poker players
Martin Zinkevich and I collaborated on this work
The CPRG’s goal: Create poker programs to beat the world’s best poker players

Martin Zinkevich and I collaborated on this work
  • This is a huge understatement
Poker is a collection of wagering card games
Poker is a collection of wagering card games

Texas Hold’em is considered to be the most strategic variant
Texas Hold’em Poker

- Poker is a collection of wagering card games
- Texas Hold’em is considered to be the most strategic variant
- Players play a series of short games against each other
Poker is a collection of wagering card games
Texas Hold’em is considered to be the most strategic variant
Players play a series of short games against each other
Goal: Win as much money as possible from opponents over this series of games
As the game progresses, more cards are revealed
Heads-Up Texas Hold’em Poker

- As the game progresses, more cards are revealed
  - Private cards that only one player can see and use
  - Public cards that all players can see and use
As the game progresses, more cards are revealed
- Private cards that only one player can see and use
- Public cards that all players can see and use

Players alternate taking actions:
- Bet: Make a wager that their cards will be the best
- Call: Match the opponent’s wager
- Fold: Surrender this game, and begin a new one.
So, why do we care about poker?

Poker is stochastic and has imperfect information...like the real world.

Exploitation is important.

Approaches for other games (such as alpha-beta) don’t apply here— we need to find new techniques.

Our techniques are applicable beyond poker.
Heads-Up Texas Hold’em Poker

- So, why do we care about poker?
So, why do we care about poker?

Poker is stochastic and has imperfect information
So, why do we care about poker?

Poker is stochastic and has imperfect information

...like the real world
Heads-Up Texas Hold’em Poker

- So, why do we care about poker?
- Poker is stochastic and has imperfect information
- ...like the real world
- Exploitation is important
So, why do we care about poker?

- Poker is stochastic and has imperfect information
- ...like the real world
- Exploitation is important
- Approaches for other games (such as alpha-beta) don’t apply here — we need to find new techniques
So, why do we care about poker?
Poker is stochastic and has imperfect information
...like the real world
Exploitation is important
Approaches for other games (such as alpha-beta) don’t apply here —
we need to find new techniques
Our techniques are applicable beyond poker
Because of hidden information, some game states are indistinguishable.
Because of hidden information, some game states are indistinguishable.

An *information set* is a set of game states that we cannot tell apart.
Because of hidden information, some game states are indistinguishable.

An *information set* is a set of game states that we cannot tell apart.

We have to play the same way for every game state in an information set.
Because of hidden information, some game states are indistinguishable.

An *information set* is a set of game states that we cannot tell apart.

We have to play the same way for every game state in an information set.

A *behavioral strategy* is a probability distribution over actions for each information set.
Poker is big — $10^{18}$ game states
• Poker is big — $10^{18}$ game states
• We abstract the cards into buckets to make the size more reasonable — $10^{12}$
- Poker is big — $10^{18}$ game states
- We abstract the cards into buckets to make the size more reasonable — $10^{12}$
- Poker strategies for the abstract game are still powerful in the “real” game, but there is a loss
1 Introduction

2 Playing to Not Lose: Counterfactual Regret Minimization

3 Playing to Win: Frequentist Best Response

4 Playing to Win, Carefully: Restricted Nash Response

5 Competition Results

6 Conclusion
First approach: a strategy that works against anyone
• First approach: a strategy that works against anyone
• Nash Equilibrium: strategy for each player, where no player can do better by unilaterally changing their strategy
First approach: a strategy that works against anyone

Nash Equilibrium: strategy for each player, where no player can do better by unilaterally changing their strategy

Approximation to a Nash equilibrium: no player can do better than ε by switching
ε-Nash equilibria

- Unbeatable (within its abstraction)
\(\epsilon\)-Nash equilibria

- Unbeatable (within its abstraction)
- The strategy can win if the opponent makes mistakes
- Nash equilibria

- Unbeatable (within its abstraction)
- The strategy can win if the opponent makes mistakes
- ...thus “playing to not lose”
$\epsilon$-Nash equilibria

- Unbeatable (within its abstraction)
- The strategy can win if the opponent makes mistakes
- ...thus “playing to not lose”
- (We still use these strategies to win)
- Unbeatable (within its abstraction)
- The strategy can win if the opponent makes mistakes
- ...thus “playing to not lose”
- (We still use these strategies to win)
- Can be found through linear programming, requires memory proportional to number of game states
\( \epsilon \)-Nash equilibria

- Unbeatable (within its abstraction)
- The strategy can win if the opponent makes mistakes
- ...thus “playing to not lose”
- (We still use these strategies to win)
- Can be found through linear programming, requires memory proportional to number of game states
- Counterfactual Regret Minimization requires memory proportional to number of information sets — much smaller.
Unbeatable (within its abstraction)
The strategy can win if the opponent makes mistakes
...thus “playing to not lose”
(We still use these strategies to win)
Can be found through linear programming, requires memory proportional to number of game states
Counterfactual Regret Minimization requires memory proportional to number of information sets — much smaller.
Poker has $3.16 \times 10^{17}$ game states and $3.19 \times 10^{14}$ information sets
Counterfactual Regret Minimization: Theory

- Play $T$ games of poker, updating your strategy on each round
Counterfactual Regret Minimization: Theory

- Play $T$ games of poker, updating your strategy on each round
- Find the best strategy you could have used for all of those games

Define Average Overall Regret as:

$$\frac{1}{T} \sum_{t=1}^{T} \left( \text{Value of best strategy} - \text{Value of your strategy} \right)$$

If we minimize Average Overall Regret, the average strategy used over the $T$ games approaches a Nash equilibrium.
Counterfactual Regret Minimization: Theory

- Play $T$ games of poker, updating your strategy on each round
- Find the best strategy you could have used for all of those games
- Define Average Overall Regret as:
  \[ \frac{1}{T} \sum_{t=1}^{T} \left( \text{(Value of best strategy)} - \text{(Value of your strategy)} \right) \]
Counterfactual Regret Minimization: Theory

- Play $T$ games of poker, updating your strategy on each round
- Find the best strategy you could have used for all of those games
- Define Average Overall Regret as:
  \[ \frac{1}{T} \sum_{t=1}^{T} \left( (\text{Value of best strategy}) - (\text{Value of your strategy}) \right) \]
- If we minimize Average Overall Regret, the *average strategy* used over the $T$ games approaches a Nash equilibrium
Counterfactual Regret Minimization: Theory

- Play $T$ games of poker, updating your strategy on each round
- Find the best strategy you could have used for all of those games
- Define Average Overall Regret as:
  \[ \frac{1}{T} \sum_{t=1}^{T} \left( (\text{Value of best strategy}) - (\text{Value of your strategy}) \right) \]
- If we minimize Average Overall Regret, the *average strategy* used over the $T$ games approaches a Nash equilibrium
- How do we minimize Average Overall Regret?
Immediate Counterfactual Regret

- Break down overall regret into the regret for each action at each information set.
Immediate Counterfactual Regret

- Break down overall regret into the regret for each action at each information set.
- Regret: How much *more* utility we could have had if we always took some action instead of using our strategy.

Our Choice

- Opponent Choice
  - Fold
  - Call
  - Raise

Our Cards:

- K♠K♣
Immediate Counterfactual Regret

- Break down overall regret into the regret for each action at each information set.
- Regret: How much *more* utility we could have had if we always took some action instead of using our strategy.
- Immediate Counterfactual Regret: Weight this regret by the probability of the opponent reaching the information set.
Immediate Counterfactual Regret

- Break down overall regret into the regret for each action at each information set
- Regret: How much *more* utility we could have had if we always took some action instead of using our strategy
- Immediate Counterfactual Regret: Weight this regret by the probability of the opponent reaching the information set
- Average Overall Regret is less than the sum of Immediate Counterfactual Regret
Immediate Counterfactual Regret

- Break down overall regret into the regret for each action at each information set
- Regret: How much *more* utility we could have had if we always took some action instead of using our strategy
- Immediate Counterfactual Regret: Weight this regret by the probability of the opponent reaching the information set
- Average Overall Regret is less than the sum of Immediate Counterfactual Regret
- So, if we can minimize our immediate counterfactual regret *at each information set*, then we approach a Nash equilibrium
Initialize the strategies’ action probabilities to a uniform distribution
Counterfactual Regret Minimization: Basic Idea

- Initialize the strategies’ action probabilities to a uniform distribution
- Repeat:
  - (General) Iterate over all chance outcomes
Initialize the strategies’ action probabilities to a uniform distribution
Repeat:
  (General) Iterate over all chance outcomes
  (Poker-specific) Deal cards to each player, as if playing the game
Initialize the strategies’ action probabilities to a uniform distribution

Repeat:

- (General) Iterate over all chance outcomes
- (Poker-specific) Deal cards to each player, as if playing the game
- Recurse over all choice nodes. Update the action probabilities at each choice node to minimize regret at that node.
Initialize the strategies’ action probabilities to a uniform distribution

Repeat:
- (General) Iterate over all chance outcomes
- (Poker-specific) Deal cards to each player, as if playing the game
- Recurse over all choice nodes. Update the action probabilities at each choice node to minimize regret at that node.

How do we update the action probabilities after each game?
Counterfactual Regret

Compute expected value of each action
Calculate the regret for not taking each action
(Regret: Difference between the EV for taking an action and the strategy's EV)
Counterfactual Regret: Regret weighted by opponent's probability of reaching this state
Add up Counterfactual Regret over all games
Assign new probabilities proportional to accumulated positive CFR
Strategy's EV: 4
Regret: (-7, 2, 5)
Total CFR: (-3.5, 1, 2.5)
New Probabilities: (0, 0.3, 0.7)
Counterfactual Regret

- Compute expected value of each action

![Diagram showing decision nodes and expected values]

- Strategy’s EV: 4

- Counterfactual Regret: Compute expected value of each action.
Counterfactual Regret

- Compute expected value of each action
- Calculate the *regret* for not taking each action
  (Regret: Difference between the EV for taking an action and the strategy’s EV)

- Strategy’s EV: 4
- Regret: (-7, 2, 5)
Counterfactual Regret

- Compute expected value of each action
- Calculate the *regret* for not taking each action
  (Regret: Difference between the EV for taking an action and the strategy’s EV)
- Counterfactual Regret: Regret weighted by opponent’s probability of reaching this state
- Add up Counterfactual Regret over all games

- Strategy’s EV: 4
- Regret: (-7, 2, 5)
- Total CFR: (-3.5, 1, 2.5)
Counterfactual Regret

- Compute expected value of each action
- Calculate the *regret* for not taking each action
  (Regret: Difference between the EV for taking an action and the strategy’s EV)
- Counterfactual Regret: Regret weighted by opponent’s probability of reaching this state
- Add up Counterfactual Regret over all games
- Assign new probabilities proportional to accumulated positive CFR

- Strategy’s EV: 4
- Regret: (-7, 2, 5)
- Total CFR: (-3.5, 1, 2.5)
- New Probabilities: (0, 0.3, 0.7)
Counterfactual Regret Example 2

Strategy's EV: -8.1
Regret: (5.1, 2.1, -0.9)
Total CFR: (1.6, 3.1, 1.6)
New Probabilities: (0.25, 0.5, 0.25)
Counterfactual Regret Example 2

- Strategy’s EV: -8.1
Counterfactual Regret Example 2

- Strategy’s EV: -8.1
- Regret: (5.1, 2.1, -0.9)
Counterfactual Regret Example 2

- Strategy’s EV: -8.1
- Regret: (5.1, 2.1, -0.9)
- Total CFR: (1.6, 3.1, 1.6)
Counterfactual Regret Example 2

- Strategy’s EV: -8.1
- Regret: (5.1, 2.1, -0.9)
- Total CFR: (1.6, 3.1, 1.6)
- New Probabilities: (0.25, 0.5, 0.25)
Counterfactual Regret Minimization approaches a Nash equilibrium - how fast does it get there?
Counterfactual Regret Minimization approaches a Nash equilibrium - how fast does it get there?

General: \# iterations grows quadratically with \# information sets

Poker: \# iterations grows linearly with \# information sets (Because seeing a few samples of the states in an information set is enough to choose a good strategy for that information set)

In practical terms: we can solve very large games (10^{12} states) in under two weeks. That's two orders of magnitude larger than was previously possible.
Counterfactual Regret Minimization approaches a Nash equilibrium - how fast does it get there?

- General: \# iterations grows quadratically with \# information sets
- Poker: \# iterations grows \textit{linearly} with \# information sets
Counterfactual Regret Minimization approaches a Nash equilibrium - how fast does it get there?

- General: \# iterations grows quadratically with \# information sets
- Poker: \# iterations grows linearly with \# information sets
- (Because seeing a few samples of the states in an information set is enough to choose a good strategy for that information set)
Counterfactual Regret Minimization approaches a Nash equilibrium - how fast does it get there?

- General: \# iterations grows quadratically with \# information sets
- Poker: \# iterations grows *linearly* with \# information sets
- (Because seeing a few samples of the states in an information set is enough to choose a good strategy for that information set)

In practical terms: we can solve very large games \((10^{12} \text{ states})\) in under two weeks

That’s two orders of magnitude larger than was previously possible
Convergence to a Nash Equilibrium

![Graph showing convergence to a Nash Equilibrium](image)

<table>
<thead>
<tr>
<th>Abstraction</th>
<th>Size (game states) ($\times 10^9$)</th>
<th>Iterations ($\times 10^6$)</th>
<th>Time (h)</th>
<th>Exp (mb/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>6.45</td>
<td>100</td>
<td>33</td>
<td>3.4</td>
</tr>
<tr>
<td>6</td>
<td>27.7</td>
<td>200</td>
<td>75</td>
<td>3.1</td>
</tr>
<tr>
<td>8</td>
<td>276</td>
<td>750</td>
<td>261</td>
<td>2.7</td>
</tr>
<tr>
<td>10</td>
<td>1646</td>
<td>2000</td>
<td>326</td>
<td>2.2</td>
</tr>
</tbody>
</table>
### Comparison to the 2006 AAAI Competition

<table>
<thead>
<tr>
<th></th>
<th>Hyperborean</th>
<th>Bluffbot</th>
<th>Monash</th>
<th>Teddy</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smallbot2298</td>
<td>61</td>
<td>113</td>
<td>695</td>
<td>474</td>
<td>336</td>
</tr>
<tr>
<td>CFR8</td>
<td>106</td>
<td>170</td>
<td>746</td>
<td>517</td>
<td>385</td>
</tr>
</tbody>
</table>
Counterfactual Regret Minimization: Conclusions

- Approaches Nash Equilibria faster and with less memory than older techniques
Counterfactual Regret Minimization: Conclusions

- Approaches Nash Equilibria faster and with less memory than older techniques
- The resulting strategies are robust — they work well against any opponent
Counterfactual Regret Minimization: Conclusions

- Approaches Nash Equilibria faster and with less memory than older techniques
- The resulting strategies are *robust* — they work well against any opponent
- But... How exploitable are the opponents?
Approaches Nash Equilibria faster and with less memory than older techniques

The resulting strategies are robust — they work well against any opponent

But... How exploitable are the opponents?

How much better could an exploitive strategy do?
Counterfactual Regret Minimization: Conclusions

- Approaches Nash Equilibria faster and with less memory than older techniques
- The resulting strategies are *robust* — they work well against any opponent
- But... How exploitable are the opponents?
- How much better could an exploitive strategy do?
- ”Playing to Not Lose”
1 Introduction

2 Playing to Not Lose: Counterfactual Regret Minimization

3 Playing to Win: Frequentist Best Response

4 Playing to Win, Carefully: Restricted Nash Response

5 Competition Results

6 Conclusion
Best Response: best possible counter-strategy to some strategy
Best Response: best possible counter-strategy to some strategy

Useful for a few reasons:
- Tells you how exploitable that strategy is
- Could use it during a match to win
Best Response Challenges

- "real" best response is intractable
Best Response Challenges

- “real” best response is intractable
- abstract game best response is easy, but has some challenges:
  - Need to actually have the opponent’s strategy
  - Resulting counter-strategy plays in the same abstraction as the strategy
Best Response Challenges

- "real" best response is intractable

abstract game best response is easy, but has some challenges:
  - Need to actually have the opponent’s strategy
  - Resulting counter-strategy plays in the same abstraction as the strategy

  (Bigger abstraction == better counter-strategy)
Motivating Frequentist Best Response

- We’d like to make best response counter-strategies with fewer restrictions:
  - What if we don’t have the actual strategy, only observations?
  - What if we want to choose the abstraction that the counter-strategy uses?
Frequentist Best Response: Basic Idea

- Observe lots of real-game data — say, 1 million hands
Frequentist Best Response: Basic Idea

- Observe lots of real-game data — say, 1 million hands
- Abstract the data, and do frequency counts on how often actions are taken in each choice node
Observe lots of real-game data — say, 1 million hands
Abstract the data, and do frequency counts on how often actions are taken in each choice node
Construct an opponent model, where action probabilities are just the action frequencies
Frequentist Best Response: Basic Idea

- Observe lots of real-game data — say, 1 million hands
- Abstract the data, and do frequency counts on how often actions are taken in each choice node
- Construct an opponent model, where action probabilities are just the action frequencies
- Find the abstract game best response to the opponent model
Frequentist Best Response: Basic Idea

- Observe lots of real-game data — say, 1 million hands
- Abstract the data, and do frequency counts on how often actions are taken in each choice node
- Construct an opponent model, where action probabilities are just the action frequencies
- Find the abstract game best response to the opponent model
- Use the counter-strategy to play against the strategy in the real game
Abstracting the data
There’s a few variables you need to get right:
There’s a few variables you need to get right:

- Who is the strategy playing against for the million hands? (Self play is bad, because it doesn’t explore the whole strategy space)
- What do you do in states you never observe? (We assume they call)
Performance of FBR Counter-strategies to Several Opponents as Training Hands Varies
### Frequentist Best Response

<table>
<thead>
<tr>
<th></th>
<th>PsOpti4</th>
<th>PsOpti6</th>
<th>Attack60</th>
<th>Attack80</th>
<th>Smallbot1239</th>
<th>Smallbot1399</th>
<th>Smallbot2298</th>
<th>CFR5</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FBR-PsOpti4</strong></td>
<td>137</td>
<td>-163</td>
<td>-227</td>
<td>-231</td>
<td>-106</td>
<td>-85</td>
<td>-144</td>
<td>-210</td>
<td>-129</td>
</tr>
<tr>
<td><strong>FBR-PsOpti6</strong></td>
<td>-79</td>
<td>330</td>
<td>-68</td>
<td>-89</td>
<td>-36</td>
<td>-23</td>
<td>-48</td>
<td>-97</td>
<td>-14</td>
</tr>
<tr>
<td><strong>FBR-Attack60</strong></td>
<td>-442</td>
<td>-499</td>
<td>2170</td>
<td>-701</td>
<td>-359</td>
<td>-305</td>
<td>-377</td>
<td>-620</td>
<td>-142</td>
</tr>
<tr>
<td><strong>FBR-Attack80</strong></td>
<td>-312</td>
<td>-281</td>
<td>-557</td>
<td>1048</td>
<td>-251</td>
<td>-231</td>
<td>-266</td>
<td>-331</td>
<td>-148</td>
</tr>
<tr>
<td><strong>FBR-Smallbot1239</strong></td>
<td>-20</td>
<td>105</td>
<td>-89</td>
<td>-42</td>
<td>106</td>
<td>91</td>
<td>-32</td>
<td>-87</td>
<td>3</td>
</tr>
<tr>
<td><strong>FBR-Smallbot1399</strong></td>
<td>-43</td>
<td>38</td>
<td>-48</td>
<td>-77</td>
<td>75</td>
<td>118</td>
<td>-46</td>
<td>-109</td>
<td>-11</td>
</tr>
<tr>
<td><strong>FBR-Smallbot2298</strong></td>
<td>-39</td>
<td>51</td>
<td>-50</td>
<td>-26</td>
<td>42</td>
<td>50</td>
<td>33</td>
<td>-41</td>
<td>2</td>
</tr>
<tr>
<td><strong>CFR5</strong></td>
<td>36</td>
<td>123</td>
<td>93</td>
<td>41</td>
<td>70</td>
<td>68</td>
<td>17</td>
<td>0</td>
<td>56</td>
</tr>
</tbody>
</table>

- **Max**: 137, 330, 2170, 1048, 106, 118, 33, 0

- **Observations**:
  - The diagonal has the matches where the counter-strategy plays against its intended opponent. These scores are all good - significantly higher than the CFR strategy does.
  - Everything off the diagonal is horrible.

- **Columns** are poker strategies we've produced in the past.
- **Rows** are counter-strategies to each strategy.
- **CFR5** is a Counterfactual Regret Minimization strategy.
### Frequentist Best Response

<table>
<thead>
<tr>
<th></th>
<th>PsOpti4</th>
<th>PsOpti6</th>
<th>Attack60</th>
<th>Attack80</th>
<th>Smallbot1239</th>
<th>Smallbot1399</th>
<th>Smallbot2298</th>
<th>CFR5</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBR-PsOpti4</td>
<td>137</td>
<td>-163</td>
<td>-227</td>
<td>-231</td>
<td>-106</td>
<td>-85</td>
<td>-144</td>
<td>-210</td>
<td>-129</td>
</tr>
<tr>
<td>FBR-PsOpti6</td>
<td>-79</td>
<td>330</td>
<td>-68</td>
<td>-89</td>
<td>-36</td>
<td>-23</td>
<td>-48</td>
<td>-97</td>
<td>-14</td>
</tr>
<tr>
<td>FBR-Attack60</td>
<td>-442</td>
<td>-499</td>
<td>2170</td>
<td>-701</td>
<td>-359</td>
<td>-305</td>
<td>-377</td>
<td>-620</td>
<td>-142</td>
</tr>
<tr>
<td>FBR-Attack80</td>
<td>-312</td>
<td>-281</td>
<td>-557</td>
<td>1048</td>
<td>-251</td>
<td>-231</td>
<td>-266</td>
<td>-331</td>
<td>-148</td>
</tr>
<tr>
<td>FBR-Smallbot1239</td>
<td>-20</td>
<td>105</td>
<td>-89</td>
<td>-42</td>
<td>106</td>
<td>91</td>
<td>-32</td>
<td>-87</td>
<td>3</td>
</tr>
<tr>
<td>FBR-Smallbot1399</td>
<td>-43</td>
<td>38</td>
<td>-48</td>
<td>-77</td>
<td>75</td>
<td>118</td>
<td>-46</td>
<td>-109</td>
<td>-11</td>
</tr>
<tr>
<td>FBR-Smallbot2298</td>
<td>-39</td>
<td>51</td>
<td>-50</td>
<td>-26</td>
<td>42</td>
<td>50</td>
<td>33</td>
<td>-41</td>
<td>2</td>
</tr>
<tr>
<td>CFR5</td>
<td>36</td>
<td>123</td>
<td>93</td>
<td>41</td>
<td>70</td>
<td>68</td>
<td>17</td>
<td>0</td>
<td>56</td>
</tr>
<tr>
<td><strong>Max</strong></td>
<td>137</td>
<td>330</td>
<td>2170</td>
<td>1048</td>
<td>106</td>
<td>118</td>
<td>33</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

- Columns are poker strategies we’ve produced in the past
- Rows are counter-strategies to each strategy
- CFR5 is a Counterfactual Regret Minimization strategy
- Two observations:
  - The diagonal has the matches where the counter-strategy plays against its intended opponent. These scores are all good - significantly higher than the CFR strategy does
  - Everything off the diagonal is horrible
Frequentist Best Response: Conclusions

"Playing to Win"
Frequentist Best Response: Conclusions

- "Playing to Win"
- Frequentist Best Response counter-strategies are useful for defeating specific opponents

However, they are brittle—when used against other opponents, even weak ones, they can lose badly.

Is there a way to keep the exploitiveness of FBR counter-strategies, while also gaining the robustness of CFR strategies?
"Playing to Win"

Frequentist Best Response counter-strategies are useful for defeating specific opponents

We also use them to evaluate our strategies, to see how weak they are
Frequentist Best Response: Conclusions

- "Playing to Win"
- Frequentist Best Response counter-strategies are useful for defeating specific opponents
- We also use them to evaluate our strategies, to see how weak they are
- However, they are brittle — when used against other opponents, even weak ones, they can lose badly.
"Playing to Win"

Frequentist Best Response counter-strategies are useful for defeating specific opponents.

We also use them to evaluate our strategies, to see how weak they are.

However, they are brittle — when used against other opponents, even weak ones, they can lose badly.

Is there a way to keep the exploitiveness of FBR counter-strategies, while also gaining the robustness of CFR strategies?
Exploiting opponents is important — we’d like to win more money than the Counterfactual Regret Minimization strategies do.
Exploiting opponents is important — we’d like to win more money than the Counterfactual Regret Minimization strategies do.

Frequentist Best Response strategies win lots of money, but are terrible against the wrong opponent.
Exploiting opponents is important — we’d like to win more money than the Counterfactual Regret Minimization strategies do.

Frequentist Best Response strategies win lots of money, but are terrible against the wrong opponent.

We’d like a compromise: a strategy that exploits an opponent (or class of opponents), but is also robust against arbitrary opponents.
We suspect our opponent will use some strategy.
We suspect our opponent will use some strategy

What if they only used it, say, 75% of the time?
We suspect our opponent will use some strategy

What if they only used it, say, 75% of the time?

The other 25% of the time, they can do anything...
Restricted Nash Response: Motivation

- We suspect our opponent will use some strategy
- What if they only used it, say, 75% of the time?
- The other 25% of the time, they can do anything...
- ...but let's assume they play a best response to whatever we do
We suspect our opponent will use some strategy.

What if they only used it, say, 75% of the time?

The other 25% of the time, they can do anything...

...but lets assume they play a best response to whatever we do.

We now have two goals: attack the 75% “weak” strategy, and defend against the 25% “adaptive” strategy.
In CFR, we had two strategies that adapt to beat each other.
In CFR, we had two strategies that adapt to beat each other.

In RNR, we have one strategy for our player, and two for our opponent.
In CFR, we had two strategies that adapt to beat each other
In RNR, we have one strategy for our player, and two for our opponent
The opponent’s static strategy is the model we get from Frequentist Best Response
In CFR, we had two strategies that adapt to beat each other.
In RNR, we have one strategy for our player, and two for our opponent.
The opponent's static strategy is the model we get from Frequentist Best Response.
We play millions of games, where our player minimizes regret when playing against both the static and adaptive opponent.
In CFR, we had two strategies that adapt to beat each other.
In RNR, we have one strategy for our player, and two for our opponent.
The opponent’s static strategy is the model we get from Frequentist Best Response.
We play millions of games, where our player minimizes regret when playing against both the static and adaptive opponent.
The adaptive opponent minimizes regret when playing against us.
“Restricted Nash Response”: our opponent is restricted to playing the static strategy some of the time.

- We approach a Nash equilibrium in this restricted game.
In the last example, we said the opponent uses the static strategy 75% of the time.
In the last example, we said the opponent uses the static strategy 75% of the time.

This is actually just a variable, $p$. 
In the last example, we said the opponent uses the static strategy 75% of the time.

This is actually just a variable, \( p \).

Interpretations of \( p \):
- How much you care about exploiting the static strategy
- How confident you are that the opponent will actually use the static strategy
If $p$ is low, then the resulting counter-strategy is more like a Nash equilibrium.

If $p$ is high, then the resulting counter-strategy is more like a best response.
Restricted Nash Response: Picking the Percentage

PsOpti4

- X-Axis: How exploitable the counter-strategy is
- Y-Axis: How much we beat the opponent
- Labels: The value of $p$ used to generate the strategy
Don’t use a Nash equilibrium - you can win a lot by giving up a tiny amount!

Don’t use a Best Response - you can save a lot by giving up a tiny amount!
### Frequentist Best Response:

<table>
<thead>
<tr>
<th>Opponents</th>
<th>PsOpti4</th>
<th>PsOpti6</th>
<th>Attack60</th>
<th>Attack80</th>
<th>Smallbot1239</th>
<th>Smallbot1399</th>
<th>Smallbot2298</th>
<th>CFR5</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNR-PsOpti4</td>
<td>137</td>
<td>-163</td>
<td>-227</td>
<td>-231</td>
<td>-106</td>
<td>-85</td>
<td>-144</td>
<td>-210</td>
<td>-129</td>
</tr>
<tr>
<td>RNR-PsOpti6</td>
<td>-79</td>
<td>330</td>
<td>-68</td>
<td>-89</td>
<td>-36</td>
<td>-23</td>
<td>-48</td>
<td>-97</td>
<td>-14</td>
</tr>
<tr>
<td>RNR-Attack60</td>
<td>-442</td>
<td>-499</td>
<td>2170</td>
<td>-701</td>
<td>-359</td>
<td>-305</td>
<td>-377</td>
<td>-620</td>
<td>-142</td>
</tr>
<tr>
<td>RNR-Attack80</td>
<td>-312</td>
<td>-281</td>
<td>-557</td>
<td>1048</td>
<td>-251</td>
<td>-231</td>
<td>-266</td>
<td>-331</td>
<td>-148</td>
</tr>
<tr>
<td>RNR-Smallbot1239</td>
<td>-20</td>
<td>105</td>
<td>-89</td>
<td>-42</td>
<td>106</td>
<td>91</td>
<td>-32</td>
<td>-87</td>
<td>3</td>
</tr>
<tr>
<td>RNR-Smallbot1399</td>
<td>-43</td>
<td>38</td>
<td>-48</td>
<td>-77</td>
<td>75</td>
<td>118</td>
<td>-46</td>
<td>-109</td>
<td>-11</td>
</tr>
<tr>
<td>RNR-Smallbot2298</td>
<td>-39</td>
<td>51</td>
<td>-50</td>
<td>-26</td>
<td>42</td>
<td>50</td>
<td>33</td>
<td>-41</td>
<td>2</td>
</tr>
<tr>
<td>CFR5</td>
<td>36</td>
<td>123</td>
<td>93</td>
<td>41</td>
<td>70</td>
<td>68</td>
<td>17</td>
<td>0</td>
<td>56</td>
</tr>
</tbody>
</table>

**Max**

| 137 | 330 | 2170 | 1048 | 106 | 118 | 33 | 0 |

### Restricted Nash Response:

<table>
<thead>
<tr>
<th>Opponents</th>
<th>PsOpti4</th>
<th>PsOpti6</th>
<th>Attack60</th>
<th>Attack80</th>
<th>Smallbot1239</th>
<th>Smallbot1399</th>
<th>Smallbot2298</th>
<th>CFR5</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNR-PsOpti4</td>
<td>85</td>
<td>112</td>
<td>39</td>
<td>9</td>
<td>63</td>
<td>61</td>
<td>-1</td>
<td>-23</td>
<td>43</td>
</tr>
<tr>
<td>RNR-PsOpti6</td>
<td>26</td>
<td>234</td>
<td>72</td>
<td>34</td>
<td>59</td>
<td>59</td>
<td>1</td>
<td>-28</td>
<td>57</td>
</tr>
<tr>
<td>RNR-Attack60</td>
<td>-17</td>
<td>63</td>
<td>582</td>
<td>-22</td>
<td>37</td>
<td>39</td>
<td>-9</td>
<td>-45</td>
<td>78</td>
</tr>
<tr>
<td>RNR-Attack80</td>
<td>-7</td>
<td>66</td>
<td>22</td>
<td>293</td>
<td>11</td>
<td>12</td>
<td>0</td>
<td>-29</td>
<td>46</td>
</tr>
<tr>
<td>RNR-Smallbot1239</td>
<td>38</td>
<td>130</td>
<td>68</td>
<td>31</td>
<td>111</td>
<td>106</td>
<td>9</td>
<td>-20</td>
<td>59</td>
</tr>
<tr>
<td>RNR-Smallbot1399</td>
<td>31</td>
<td>136</td>
<td>66</td>
<td>29</td>
<td>105</td>
<td>112</td>
<td>6</td>
<td>-24</td>
<td>58</td>
</tr>
<tr>
<td>RNR-Smallbot2298</td>
<td>21</td>
<td>137</td>
<td>72</td>
<td>30</td>
<td>77</td>
<td>76</td>
<td>31</td>
<td>-11</td>
<td>54</td>
</tr>
<tr>
<td>CFR5</td>
<td>36</td>
<td>123</td>
<td>93</td>
<td>41</td>
<td>70</td>
<td>68</td>
<td>17</td>
<td>0</td>
<td>56</td>
</tr>
</tbody>
</table>

**Max**

| 85 | 234 | 582 | 293 | 111 | 112 | 31 | 0 |
“Playing to Win, Carefully”
“Playing to Win, Carefully”

Restricted Nash Response makes *robust* counter-strategies
“Playing to Win, Carefully”

- Restricted Nash Response makes *robust* counter-strategies
- Exploits one opponent, minimizes weakness against all others
“Playing to Win, Carefully”

- Restricted Nash Response makes robust counter-strategies
- Exploits one opponent, minimizes weakness against all others
- If you ever have to compute a best response offline, you can do this instead. It’s not so bad if you’re right, and a life saver if you’re wrong.
Introduction

Playing to Not Lose: Counterfactual Regret Minimization

Playing to Win: Frequentist Best Response

Playing to Win, Carefully: Restricted Nash Response

Competition Results

Conclusion
We competed in two competitions at AAAI this year:
We competed in two competitions at AAAI this year:

- **Second AAAI Computer Poker Competition**
  - 3 events, 15 competitors, 43 bots
  - Used CFR strategies to get a 1st, a 2nd, and a 3rd
We competed in two competitions at AAAI this year:

- **Second AAAI Computer Poker Competition**
  - 3 events, 15 competitors, 43 bots
  - Used CFR strategies to get a 1st, a 2nd, and a 3rd

- **First Man-Machine Poker Championship**
  - Played against two poker pros, Phil Laak and Ali Eslami
  - Used CFR and RNR strategies to win one, tie one, and lose two
  - Post-game analysis suggests a different result
Introduction

Playing to Not Lose: Counterfactual Regret Minimization

Playing to Win: Frequentist Best Response

Playing to Win, Carefully: Restricted Nash Response

Competition Results

Conclusion
3 new techniques for stochastic, imperfect information games:

- **Counterfactual Regret Minimization**
  - "Playing to Not Lose"

- **Frequentist Best Response**
  - "Playing to Win"

- **Restricted Nash Response**
  - "Playing to Win, Carefully"

We proved the value of these techniques through competitive play.
3 new techniques for stochastic, imperfect information games:

- **Counterfactual Regret Minimization**
  - "Playing to Not Lose"
  - Approximate Nash Equilibrium strategies
  - Runs faster and with lower memory requirements than past techniques

- **Frequentist Best Response**
  - "Playing to Win"

- **Restricted Nash Response**
  - "Playing to Win, Carefully"
3 new techniques for stochastic, imperfect information games:

- **Counterfactual Regret Minimization**
  - "Playing to Not Lose"
  - Approximate Nash Equilibrium strategies
  - Runs faster and with lower memory requirements than past techniques

- **Frequentist Best Response**
  - "Playing to Win"
  - Finds exploitive counter-strategies for specific opponents
  - Useful for finding maximum exploitability of an opponent
  - Brittle when used against other opponents

- **Restricted Nash Response**
  - "Playing to Win, Carefully"
3 new techniques for stochastic, imperfect information games:

- **Counterfactual Regret Minimization**
  - ”Playing to Not Lose”
  - Approximate Nash Equilibrium strategies
  - Runs faster and with lower memory requirements than past techniques

- **Frequentist Best Response**
  - ”Playing to Win”
  - Finds exploitive counter-strategies for specific opponents
  - Useful for finding maximum exploitability of an opponent
  - Brittle when used against other opponents

- **Restricted Nash Response**
  - ”Playing to Win, Carefully”
  - Finds *robust* counter-strategies for specific opponents
  - Useful for exploiting a suspected tendency
  - Robust when used against other opponents
3 new techniques for stochastic, imperfect information games:

- **Counterfactual Regret Minimization**
  - "Playing to Not Lose"
  - Approximate Nash Equilibrium strategies
  - Runs faster and with lower memory requirements than past techniques

- **Frequentist Best Response**
  - "Playing to Win"
  - Finds exploitive counter-strategies for specific opponents
  - Useful for finding maximum exploitability of an opponent
  - Brittle when used against other opponents

- **Restricted Nash Response**
  - "Playing to Win, Carefully"
  - Finds robust counter-strategies for specific opponents
  - Useful for exploiting a suspected tendency
  - Robust when used against other opponents

- We proved the value of these techniques through competitive play
Concluding Thoughts

- There’s another Computer Poker Competition next year, and we’re hoping for another Man-Machine match
There’s another Computer Poker Competition next year, and we’re hoping for another Man-Machine match.

We have many directions to take this work:
- Better ways to manage a team of strategies
- Counter-strategies that exploit a wide variety of opponents
- ...and many other parts of the problem
Concluding Thoughts

- There’s another Computer Poker Competition next year, and we’re hoping for another Man-Machine match.
- We have many directions to take this work:
  - Better ways to manage a team of strategies
  - Counter-strategies that exploit a wide variety of opponents
  - ...and many other parts of the problem
- The CFR and RNR techniques described in this thesis are iterative:
  - The longer you run the program, the better they get
  - Over the next year, we can produce much stronger poker programs
  - The quality of human play will not improve much this year
Concluding Thoughts

- There’s another Computer Poker Competition next year, and we’re hoping for another Man-Machine match.
- We have many directions to take this work:
  - Better ways to manage a team of strategies
  - Counter-strategies that exploit a wide variety of opponents
  - ...and many other parts of the problem
- The CFR and RNR techniques described in this thesis are iterative:
  - The longer you run the program, the better they get
  - Over the next year, we can produce much stronger poker programs
  - The quality of human play will not improve much this year
- The next Man-Machine match might have a different outcome!
Questions?
Second year it’s been run
Second year it’s been run
Last year: 2 events, 5 competitors, 5 bots
This year: 3 events, 15 competitors, 43 bots
Second year it’s been run

Last year: 2 events, 5 competitors, 5 bots

This year: 3 events, 15 competitors, 43 bots

3 Events:

- Heads-Up Limit Equilibrium
- Heads-Up Limit Online Learning
- Heads-Up No-Limit
Winner determined by total matches (not dollars!) won

<table>
<thead>
<tr>
<th>Hyperborean07EQ</th>
<th>IanBot</th>
<th>GS3</th>
<th>PokeMinn</th>
<th>Quick</th>
<th>Gomel-2</th>
<th>DumboEQ</th>
<th>DumboEQ-2</th>
<th>Sequel</th>
<th>Sequel-2</th>
<th>PokeMinn-2</th>
<th>UNCC</th>
<th>Gomel</th>
<th>LeRenard</th>
<th>MonashBPP</th>
<th>MilanoEQ</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperborean07EQ</td>
<td></td>
<td>21</td>
<td>32</td>
<td>136</td>
<td>115</td>
<td>110</td>
<td>193</td>
<td>182</td>
<td>165</td>
<td>166</td>
<td>131</td>
<td>115</td>
<td>138</td>
<td>465</td>
<td>428</td>
<td>194</td>
</tr>
<tr>
<td>IanBot</td>
<td>-21</td>
<td></td>
<td></td>
<td>130</td>
<td>99</td>
<td>85</td>
<td>142</td>
<td>119</td>
<td>131</td>
<td>140</td>
<td>142</td>
<td>472</td>
<td>88</td>
<td>130</td>
<td>398</td>
<td>164</td>
</tr>
<tr>
<td>GS3</td>
<td>-32</td>
<td>-4</td>
<td></td>
<td>150</td>
<td>73</td>
<td>112</td>
<td>160</td>
<td>149</td>
<td>140</td>
<td>148</td>
<td>154</td>
<td>467</td>
<td>107</td>
<td>142</td>
<td>445</td>
<td>175</td>
</tr>
<tr>
<td>PokeMinn</td>
<td>-136</td>
<td>-130</td>
<td>-150</td>
<td>150</td>
<td>144</td>
<td>140</td>
<td>80</td>
<td>114</td>
<td>-33</td>
<td>-22</td>
<td>-24</td>
<td>373</td>
<td>265</td>
<td>127</td>
<td>627</td>
<td>421</td>
</tr>
<tr>
<td>Quick</td>
<td>-115</td>
<td>-99</td>
<td>-73</td>
<td>136</td>
<td>145</td>
<td>125</td>
<td>150</td>
<td>135</td>
<td>129</td>
<td>134</td>
<td>134</td>
<td>298</td>
<td>149</td>
<td>15</td>
<td>564</td>
<td>489</td>
</tr>
<tr>
<td>Gomel-2</td>
<td>-110</td>
<td>-85</td>
<td>-112</td>
<td>140</td>
<td>-19</td>
<td>140</td>
<td>206</td>
<td>200</td>
<td>135</td>
<td>150</td>
<td>16</td>
<td>275</td>
<td>232</td>
<td>136</td>
<td>802</td>
<td>859</td>
</tr>
<tr>
<td>DumboEQ</td>
<td>-193</td>
<td>-142</td>
<td>-160</td>
<td>200</td>
<td>-235</td>
<td>-206</td>
<td>133</td>
<td>135</td>
<td>129</td>
<td>134</td>
<td>139</td>
<td>230</td>
<td>149</td>
<td>15</td>
<td>564</td>
<td>489</td>
</tr>
<tr>
<td>DumboEQ-2</td>
<td>-182</td>
<td>-119</td>
<td>-149</td>
<td>150</td>
<td>-135</td>
<td>-200</td>
<td>-133</td>
<td>136</td>
<td>87</td>
<td>82</td>
<td>83</td>
<td>52</td>
<td>211</td>
<td>54</td>
<td>808</td>
<td>762</td>
</tr>
<tr>
<td>Sequel</td>
<td>-165</td>
<td>-131</td>
<td>-140</td>
<td>100</td>
<td>-125</td>
<td>-135</td>
<td>67</td>
<td>87</td>
<td>19</td>
<td>130</td>
<td>167</td>
<td>-17</td>
<td>141</td>
<td>92</td>
<td>556</td>
<td>556</td>
</tr>
<tr>
<td>Sequel-2</td>
<td>-166</td>
<td>-140</td>
<td>-148</td>
<td>90</td>
<td>-121</td>
<td>-150</td>
<td>-64</td>
<td>-82</td>
<td>-19</td>
<td>125</td>
<td>-17</td>
<td>34</td>
<td>141</td>
<td>74</td>
<td>583</td>
<td>526</td>
</tr>
<tr>
<td>PokeMinn-2</td>
<td>-131</td>
<td>-142</td>
<td>-154</td>
<td>40</td>
<td>-134</td>
<td>-16</td>
<td>-55</td>
<td>-83</td>
<td>-130</td>
<td>-125</td>
<td>96</td>
<td>123</td>
<td>60</td>
<td>770</td>
<td>748</td>
<td>57</td>
</tr>
</tbody>
</table>
**Winner determined by total matches (not dollars!) won**

- Emphasizes winning, not exploiting

<table>
<thead>
<tr>
<th></th>
<th>Hyperborean07EQ</th>
<th>IanBot</th>
<th>GS3</th>
<th>PokeMinn</th>
<th>Quick</th>
<th>Gomel-2</th>
<th>DumboEQ</th>
<th>DumboEQ-2</th>
<th>Sequel</th>
<th>Sequel-2</th>
<th>PokeMinn-2</th>
<th>UNCC</th>
<th>Gomel</th>
<th>LeRenard</th>
<th>MonashBPP</th>
<th>MilanoEQ</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperborean07EQ</td>
<td>21</td>
<td>32</td>
<td>136</td>
<td>115</td>
<td>110</td>
<td>193</td>
<td>182</td>
<td>165</td>
<td>166</td>
<td>131</td>
<td>454</td>
<td>115</td>
<td>138</td>
<td>465</td>
<td>428</td>
<td>194</td>
<td></td>
</tr>
<tr>
<td>IanBot</td>
<td>-21</td>
<td>4</td>
<td>130</td>
<td>99</td>
<td>85</td>
<td>142</td>
<td>119</td>
<td>131</td>
<td>140</td>
<td>142</td>
<td>472</td>
<td>88</td>
<td>130</td>
<td>408</td>
<td>398</td>
<td>164</td>
<td></td>
</tr>
<tr>
<td>GS3</td>
<td>-32</td>
<td>-4</td>
<td>150</td>
<td>73</td>
<td>112</td>
<td>160</td>
<td>149</td>
<td>140</td>
<td>148</td>
<td>154</td>
<td>467</td>
<td>107</td>
<td>142</td>
<td>412</td>
<td>445</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>PokeMinn</td>
<td>-136</td>
<td>-130</td>
<td>-150</td>
<td>40</td>
<td>144</td>
<td>80</td>
<td>76</td>
<td>-33</td>
<td>-22</td>
<td>-24</td>
<td>373</td>
<td>265</td>
<td>127</td>
<td>627</td>
<td>421</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>Quick</td>
<td>-115</td>
<td>-99</td>
<td>-73</td>
<td>-40</td>
<td>19</td>
<td>235</td>
<td>135</td>
<td>125</td>
<td>121</td>
<td>134</td>
<td>298</td>
<td>149</td>
<td>15</td>
<td>564</td>
<td>489</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>Gomel-2</td>
<td>-110</td>
<td>-85</td>
<td>-112</td>
<td>-144</td>
<td>-19</td>
<td>206</td>
<td>200</td>
<td>135</td>
<td>150</td>
<td>16</td>
<td>275</td>
<td>232</td>
<td>136</td>
<td>802</td>
<td>859</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>DumboEQ</td>
<td>-193</td>
<td>-142</td>
<td>-160</td>
<td>-80</td>
<td>-235</td>
<td>-206</td>
<td>133</td>
<td>67</td>
<td>64</td>
<td>55</td>
<td>23</td>
<td>300</td>
<td>13</td>
<td>774</td>
<td>672</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>DumboEQ-2</td>
<td>-182</td>
<td>-119</td>
<td>-149</td>
<td>-76</td>
<td>-135</td>
<td>-200</td>
<td>-133</td>
<td>87</td>
<td>82</td>
<td>83</td>
<td>-52</td>
<td>271</td>
<td>54</td>
<td>808</td>
<td>762</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Sequel</td>
<td>-165</td>
<td>-131</td>
<td>-140</td>
<td>-33</td>
<td>-125</td>
<td>-135</td>
<td>-67</td>
<td>-87</td>
<td>19</td>
<td>130</td>
<td>167</td>
<td>-17</td>
<td>92</td>
<td>556</td>
<td>556</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>Sequel-2</td>
<td>-166</td>
<td>-140</td>
<td>-148</td>
<td>-22</td>
<td>-121</td>
<td>-150</td>
<td>-64</td>
<td>-82</td>
<td>-19</td>
<td>125</td>
<td>174</td>
<td>-4</td>
<td>74</td>
<td>583</td>
<td>526</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>PokeMinn-2</td>
<td>-131</td>
<td>-142</td>
<td>-154</td>
<td>-24</td>
<td>-134</td>
<td>-16</td>
<td>-55</td>
<td>-83</td>
<td>-130</td>
<td>-125</td>
<td>96</td>
<td>123</td>
<td>60</td>
<td>770</td>
<td>748</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>Gomel</td>
<td>-115</td>
<td>-88</td>
<td>-107</td>
<td>-265</td>
<td>-149</td>
<td>-232</td>
<td>-300</td>
<td>-271</td>
<td>17</td>
<td>4</td>
<td>-123</td>
<td>-95</td>
<td>96</td>
<td>779</td>
<td>993</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>MonashBPP</td>
<td>-465</td>
<td>-408</td>
<td>-412</td>
<td>-627</td>
<td>-564</td>
<td>-802</td>
<td>-774</td>
<td>-808</td>
<td>-556</td>
<td>-583</td>
<td>-770</td>
<td>-553</td>
<td>-779</td>
<td>489</td>
<td>489</td>
<td>-539</td>
<td></td>
</tr>
</tbody>
</table>
Winner determined by total matches (not dollars!) won
Emphasizes winning, not exploiting
Took first place, using a CFR bot

<table>
<thead>
<tr>
<th></th>
<th>Hyperborean07EQ</th>
<th>IanBot</th>
<th>GS3</th>
<th>PokeMinn</th>
<th>Quick</th>
<th>Gomel-2</th>
<th>DumboEQ</th>
<th>DumboEQ-2</th>
<th>sequel</th>
<th>sequel-2</th>
<th>PokeMinn-2</th>
<th>UNCC</th>
<th>Gomel</th>
<th>LeRenard</th>
<th>MonashBPP</th>
<th>MilanoEQ</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperborean07EQ</td>
<td>-32</td>
<td>-4</td>
<td>136</td>
<td>115</td>
<td>110</td>
<td>193</td>
<td>182</td>
<td>165</td>
<td>166</td>
<td>166</td>
<td>131</td>
<td>454</td>
<td>115</td>
<td>138</td>
<td>465</td>
<td>428</td>
<td>194</td>
</tr>
<tr>
<td>IanBot</td>
<td>-21</td>
<td>4</td>
<td>130</td>
<td>99</td>
<td>85</td>
<td>142</td>
<td>119</td>
<td>131</td>
<td>140</td>
<td>142</td>
<td>142</td>
<td>472</td>
<td>88</td>
<td>130</td>
<td>408</td>
<td>398</td>
<td>164</td>
</tr>
<tr>
<td>GS3</td>
<td>-136</td>
<td>-130</td>
<td>-150</td>
<td>40</td>
<td>144</td>
<td>235</td>
<td>135</td>
<td>125</td>
<td>121</td>
<td>134</td>
<td>373</td>
<td>265</td>
<td>127</td>
<td>627</td>
<td>421</td>
<td>111</td>
<td>175</td>
</tr>
<tr>
<td>PokeMinn</td>
<td>-115</td>
<td>-99</td>
<td>-73</td>
<td>-40</td>
<td>-19</td>
<td>235</td>
<td>135</td>
<td>125</td>
<td>121</td>
<td>134</td>
<td>373</td>
<td>265</td>
<td>127</td>
<td>627</td>
<td>421</td>
<td>111</td>
<td>175</td>
</tr>
<tr>
<td>Quick</td>
<td>-110</td>
<td>-85</td>
<td>-112</td>
<td>-144</td>
<td>-19</td>
<td>206</td>
<td>200</td>
<td>135</td>
<td>150</td>
<td>166</td>
<td>130</td>
<td>167</td>
<td>-17</td>
<td>92</td>
<td>556</td>
<td>556</td>
<td>46</td>
</tr>
<tr>
<td>Gomel-2</td>
<td>-193</td>
<td>-142</td>
<td>-160</td>
<td>-80</td>
<td>-235</td>
<td>206</td>
<td>200</td>
<td>135</td>
<td>150</td>
<td>166</td>
<td>130</td>
<td>167</td>
<td>-17</td>
<td>92</td>
<td>556</td>
<td>556</td>
<td>46</td>
</tr>
<tr>
<td>DumboEQ</td>
<td>-182</td>
<td>-119</td>
<td>-149</td>
<td>-76</td>
<td>-135</td>
<td>-133</td>
<td>-87</td>
<td>87</td>
<td>19</td>
<td>32</td>
<td>125</td>
<td>174</td>
<td>-4</td>
<td>74</td>
<td>583</td>
<td>520</td>
<td>41</td>
</tr>
<tr>
<td>DumboEQ-2</td>
<td>-165</td>
<td>-131</td>
<td>-140</td>
<td>-33</td>
<td>-125</td>
<td>-133</td>
<td>-87</td>
<td>87</td>
<td>19</td>
<td>32</td>
<td>125</td>
<td>174</td>
<td>-4</td>
<td>74</td>
<td>583</td>
<td>520</td>
<td>41</td>
</tr>
<tr>
<td>Sequel</td>
<td>-166</td>
<td>-140</td>
<td>-148</td>
<td>-64</td>
<td>-82</td>
<td>-67</td>
<td>-87</td>
<td>19</td>
<td>32</td>
<td>32</td>
<td>125</td>
<td>174</td>
<td>-4</td>
<td>74</td>
<td>583</td>
<td>520</td>
<td>41</td>
</tr>
<tr>
<td>Sequel-2</td>
<td>-131</td>
<td>-142</td>
<td>-154</td>
<td>-64</td>
<td>-82</td>
<td>-67</td>
<td>-87</td>
<td>19</td>
<td>32</td>
<td>32</td>
<td>125</td>
<td>174</td>
<td>-4</td>
<td>74</td>
<td>583</td>
<td>520</td>
<td>41</td>
</tr>
<tr>
<td>PokeMinn-2</td>
<td>-131</td>
<td>-142</td>
<td>-154</td>
<td>-64</td>
<td>-82</td>
<td>-67</td>
<td>-87</td>
<td>19</td>
<td>32</td>
<td>32</td>
<td>125</td>
<td>174</td>
<td>-4</td>
<td>74</td>
<td>583</td>
<td>520</td>
<td>41</td>
</tr>
<tr>
<td>UNCC</td>
<td>-454</td>
<td>-472</td>
<td>-467</td>
<td>-373</td>
<td>-298</td>
<td>-55</td>
<td>-83</td>
<td>-130</td>
<td>-125</td>
<td>-130</td>
<td>96</td>
<td>123</td>
<td>60</td>
<td>770</td>
<td>748</td>
<td>57</td>
<td>57</td>
</tr>
<tr>
<td>Gomel</td>
<td>-115</td>
<td>-88</td>
<td>-107</td>
<td>-265</td>
<td>-149</td>
<td>-300</td>
<td>-271</td>
<td>17</td>
<td>4</td>
<td>-123</td>
<td>96</td>
<td>95</td>
<td>56</td>
<td>779</td>
<td>993</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Mike Johanson () Robust Strategies and Counter-Strategies November 20, 2012 51 / 65
Winner determined by total winnings (in dollars)

<table>
<thead>
<tr>
<th></th>
<th>Hyperborean07OL-2</th>
<th>Hyperborean07OL</th>
<th>GS3</th>
<th>IanBot</th>
<th>Quick</th>
<th>Gomel-2</th>
<th>PokeMinn</th>
<th>Sequel</th>
<th>Sequel-2</th>
<th>LeRenard</th>
<th>DumboOL-2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperborean07OL-2</td>
<td></td>
<td>-37</td>
<td>-27</td>
<td>-37</td>
<td>138</td>
<td>155</td>
<td>172</td>
<td>166</td>
<td>178</td>
<td>170</td>
<td>259</td>
<td>114</td>
</tr>
<tr>
<td>Hyperborean07OL</td>
<td>37</td>
<td></td>
<td>21</td>
<td>27</td>
<td>116</td>
<td>108</td>
<td>141</td>
<td>153</td>
<td>175</td>
<td>132</td>
<td>207</td>
<td>112</td>
</tr>
<tr>
<td>GS3</td>
<td>27</td>
<td>-21</td>
<td>6</td>
<td>73</td>
<td>112</td>
<td>150</td>
<td>140</td>
<td>148</td>
<td>142</td>
<td>199</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>IanBot</td>
<td>-138</td>
<td>-116</td>
<td>-73</td>
<td>-99</td>
<td>19</td>
<td>-40</td>
<td>125</td>
<td>121</td>
<td>15</td>
<td>129</td>
<td>-6</td>
<td></td>
</tr>
<tr>
<td>Gomel-2</td>
<td>-144</td>
<td>144</td>
<td>40</td>
<td>144</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-8</td>
</tr>
<tr>
<td>PokeMinn</td>
<td>-172</td>
<td>-141</td>
<td>-150</td>
<td>-130</td>
<td>40</td>
<td>144</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-35</td>
</tr>
<tr>
<td>Sequel</td>
<td></td>
<td></td>
<td>-166</td>
<td>-153</td>
<td>-140</td>
<td>-131</td>
<td>-125</td>
<td>-135</td>
<td>33</td>
<td>19</td>
<td>92</td>
<td>-1</td>
</tr>
<tr>
<td>Sequel-2</td>
<td>-178</td>
<td>-175</td>
<td>-148</td>
<td>-140</td>
<td>-121</td>
<td>-150</td>
<td>22</td>
<td>-19</td>
<td>74</td>
<td>17</td>
<td>82</td>
<td>-71</td>
</tr>
<tr>
<td>DumboOL-2</td>
<td></td>
<td></td>
<td>-259</td>
<td>-207</td>
<td>-199</td>
<td>-157</td>
<td>-129</td>
<td>-123</td>
<td>15</td>
<td>-17</td>
<td>-21</td>
<td>-110</td>
</tr>
</tbody>
</table>
- Winner determined by total winnings (in dollars)
- Took second place with a CFR bot. We just barely lost to...

<table>
<thead>
<tr>
<th></th>
<th>Hyperborean07OL-2</th>
<th>Hyperborean07OL</th>
<th>GS3</th>
<th>IanBot</th>
<th>Quick</th>
<th>Gomel-2</th>
<th>PokeMinn</th>
<th>Sequel</th>
<th>Sequel-2</th>
<th>LeRenard</th>
<th>DumboOL-2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperborean07OL-2</td>
<td>-37</td>
<td>-27</td>
<td>-37</td>
<td>138</td>
<td>155</td>
<td>172</td>
<td>166</td>
<td>178</td>
<td>170</td>
<td>259</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>Hyperborean07OL</td>
<td>37</td>
<td>27</td>
<td>21</td>
<td>116</td>
<td>108</td>
<td>141</td>
<td>153</td>
<td>175</td>
<td>132</td>
<td>207</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>GS3</td>
<td>27</td>
<td>-21</td>
<td>-6</td>
<td>73</td>
<td>112</td>
<td>150</td>
<td>140</td>
<td>148</td>
<td>142</td>
<td>199</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>IanBot</td>
<td>37</td>
<td>-27</td>
<td>-6</td>
<td>99</td>
<td>85</td>
<td>130</td>
<td>131</td>
<td>140</td>
<td>130</td>
<td>157</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>Quick</td>
<td>-138</td>
<td>-116</td>
<td>-73</td>
<td>-99</td>
<td>19</td>
<td>-40</td>
<td>125</td>
<td>121</td>
<td>15</td>
<td>129</td>
<td>-6</td>
<td></td>
</tr>
<tr>
<td>PokeMinn</td>
<td>-172</td>
<td>-141</td>
<td>-150</td>
<td>40</td>
<td>144</td>
<td>-33</td>
<td>-22</td>
<td>127</td>
<td>-15</td>
<td>-35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequel</td>
<td>-166</td>
<td>-153</td>
<td>-140</td>
<td>-131</td>
<td>-125</td>
<td>-135</td>
<td>33</td>
<td>19</td>
<td>92</td>
<td>-1</td>
<td>-71</td>
<td></td>
</tr>
<tr>
<td>Sequel-2</td>
<td>-178</td>
<td>-175</td>
<td>-148</td>
<td>-140</td>
<td>-121</td>
<td>-150</td>
<td>22</td>
<td>-19</td>
<td>74</td>
<td>17</td>
<td>-82</td>
<td></td>
</tr>
<tr>
<td>DumboOL-2</td>
<td>-259</td>
<td>-207</td>
<td>-199</td>
<td>-157</td>
<td>-129</td>
<td>-123</td>
<td>15</td>
<td>1</td>
<td>-17</td>
<td>-21</td>
<td>-110</td>
<td></td>
</tr>
</tbody>
</table>
Winner determined by total winnings (in dollars)
Took second place with a CFR bot. We just barely lost to...
...the other U of A bot
- Winner determined by total winnings (in dollars)
- Took second place with a CFR bot. We just barely lost to...
- ...the other U of A bot
- (Darse Billings and Morgan Kan)

<table>
<thead>
<tr>
<th></th>
<th>Hyperborean07OL-2</th>
<th>Hyperborean07OL</th>
<th>GS3</th>
<th>IanBot</th>
<th>Quick</th>
<th>Gomel-2</th>
<th>PokeMinn</th>
<th>Sequel</th>
<th>Sequel-2</th>
<th>LeRenard</th>
<th>DumboOL-2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperborean07OL-2</td>
<td></td>
<td></td>
<td></td>
<td>-37</td>
<td>-27</td>
<td>138</td>
<td>155</td>
<td>172</td>
<td>166</td>
<td>178</td>
<td>170</td>
<td>259</td>
</tr>
<tr>
<td>Hyperborean07OL</td>
<td>37</td>
<td></td>
<td></td>
<td>21</td>
<td>27</td>
<td>116</td>
<td>108</td>
<td>141</td>
<td>153</td>
<td>175</td>
<td>132</td>
<td>207</td>
</tr>
<tr>
<td>GS3</td>
<td>27</td>
<td>-21</td>
<td></td>
<td>6</td>
<td>73</td>
<td>112</td>
<td>150</td>
<td>140</td>
<td>148</td>
<td>142</td>
<td>199</td>
<td>98</td>
</tr>
<tr>
<td>IanBot</td>
<td>37</td>
<td>-27</td>
<td></td>
<td>-6</td>
<td>99</td>
<td>85</td>
<td>130</td>
<td>131</td>
<td>140</td>
<td>130</td>
<td>157</td>
<td>87</td>
</tr>
<tr>
<td>Quick</td>
<td>-138</td>
<td>-116</td>
<td></td>
<td>-73</td>
<td>-99</td>
<td>19</td>
<td>-40</td>
<td>125</td>
<td>121</td>
<td>15</td>
<td>129</td>
<td>-6</td>
</tr>
<tr>
<td>PokeMinn</td>
<td>-172</td>
<td>-141</td>
<td></td>
<td>-150</td>
<td>-130</td>
<td>40</td>
<td>144</td>
<td>-33</td>
<td>-22</td>
<td>127</td>
<td>-15</td>
<td>-35</td>
</tr>
<tr>
<td>Sequal</td>
<td>-166</td>
<td>-153</td>
<td></td>
<td>-140</td>
<td>-131</td>
<td>-125</td>
<td>-135</td>
<td>33</td>
<td>19</td>
<td>92</td>
<td>-1</td>
<td>-71</td>
</tr>
<tr>
<td>Sequal-2</td>
<td>-178</td>
<td>-175</td>
<td></td>
<td>-148</td>
<td>-140</td>
<td>-121</td>
<td>-150</td>
<td>22</td>
<td>-19</td>
<td>74</td>
<td>17</td>
<td>-82</td>
</tr>
<tr>
<td>DumboOL-2</td>
<td>-259</td>
<td>-207</td>
<td></td>
<td>-199</td>
<td>-157</td>
<td>-129</td>
<td>-123</td>
<td>15</td>
<td>1</td>
<td>-17</td>
<td>-21</td>
<td>-110</td>
</tr>
</tbody>
</table>
**AAAI: Heads-Up No-Limit**

- No-Limit is what you see on TV - bets can be any size

<table>
<thead>
<tr>
<th></th>
<th>BluffBot20</th>
<th>GS3</th>
<th>Hyperborean07</th>
<th>SlideRule</th>
<th>Gomel</th>
<th>Gomel-2</th>
<th>Milano</th>
<th>Manitoba</th>
<th>PokeMinn</th>
<th>Manitoba-2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>BluffBot20</td>
<td>267</td>
<td>380</td>
<td>576</td>
<td>2093</td>
<td>2885</td>
<td>3437</td>
<td>475</td>
<td>1848</td>
<td>2471</td>
<td>1603</td>
<td></td>
</tr>
<tr>
<td>GS3</td>
<td>-267</td>
<td>113</td>
<td>503</td>
<td>3161</td>
<td>124</td>
<td>1875</td>
<td>4204</td>
<td>-42055</td>
<td>5016</td>
<td>-3036</td>
<td></td>
</tr>
<tr>
<td>Hyperborean07</td>
<td>-380</td>
<td>-113</td>
<td>-48</td>
<td>6657</td>
<td>5455</td>
<td>6795</td>
<td>8697</td>
<td>12051</td>
<td>22116</td>
<td>6803</td>
<td></td>
</tr>
<tr>
<td>SlideRule</td>
<td>-576</td>
<td>-503</td>
<td>48</td>
<td>11596</td>
<td>9730</td>
<td>10337</td>
<td>10387</td>
<td>15637</td>
<td>10791</td>
<td>7494</td>
<td></td>
</tr>
<tr>
<td>Gomel</td>
<td>-2093</td>
<td>-3161</td>
<td>-6657</td>
<td>-11596</td>
<td>3184</td>
<td>8372</td>
<td>11450</td>
<td>62389</td>
<td>52325</td>
<td>12690</td>
<td></td>
</tr>
<tr>
<td>Gomel-2</td>
<td>-2885</td>
<td>-124</td>
<td>-5455</td>
<td>-9730</td>
<td>-3184</td>
<td>15078</td>
<td>11907</td>
<td>58985</td>
<td>40256</td>
<td>11650</td>
<td></td>
</tr>
<tr>
<td>Milano</td>
<td>-3437</td>
<td>-1875</td>
<td>-6795</td>
<td>-10337</td>
<td>-8372</td>
<td>-15078</td>
<td>5741</td>
<td>12719</td>
<td>27040</td>
<td>-44</td>
<td></td>
</tr>
<tr>
<td>Manitoba</td>
<td>-475</td>
<td>-4204</td>
<td>-8697</td>
<td>-10387</td>
<td>-11450</td>
<td>-11907</td>
<td>-5741</td>
<td>18817</td>
<td>50677</td>
<td>1848</td>
<td></td>
</tr>
<tr>
<td>PokeMinn</td>
<td>-1848</td>
<td>42055</td>
<td>-14051</td>
<td>-15637</td>
<td>-62389</td>
<td>-58985</td>
<td>-12719</td>
<td>-18817</td>
<td>34299</td>
<td>-12010</td>
<td></td>
</tr>
<tr>
<td>Manitoba-2</td>
<td>-2471</td>
<td>-5016</td>
<td>-22116</td>
<td>-10791</td>
<td>-52325</td>
<td>-40256</td>
<td>-27040</td>
<td>-50677</td>
<td>-34299</td>
<td>-27221</td>
<td></td>
</tr>
</tbody>
</table>

Mike Johanson () Robust Strategies and Counter-Strategies November 20, 2012 53 / 65
AAAI: Heads-Up No-Limit

- No-Limit is what you see on TV - bets can be any size
- This was our first time making a No-Limit bot

<table>
<thead>
<tr>
<th></th>
<th>BluffBot20</th>
<th>GS3</th>
<th>Hyperborean07</th>
<th>SlideRule</th>
<th>Gomel</th>
<th>Gomel-2</th>
<th>Milano</th>
<th>Manitoba</th>
<th>PokeMinn</th>
<th>Manitoba-2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>BluffBot20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS3</td>
<td>-267</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyperborean07</td>
<td>-380</td>
<td>-113</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SlideRule</td>
<td>-576</td>
<td>-503</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gomel</td>
<td>-2093</td>
<td>-3161</td>
<td>-6657</td>
<td>-11596</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gomel-2</td>
<td>-2885</td>
<td>-124</td>
<td>-5455</td>
<td>-9730</td>
<td>-3184</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milano</td>
<td>-3437</td>
<td>-1875</td>
<td>-6795</td>
<td>-10337</td>
<td>-8372</td>
<td>-15078</td>
<td></td>
<td></td>
<td>5741</td>
<td>12719</td>
<td>27040</td>
</tr>
<tr>
<td>Manitoba</td>
<td>-475</td>
<td>-4204</td>
<td>-8697</td>
<td>-10387</td>
<td>-11450</td>
<td>-11907</td>
<td>-5741</td>
<td></td>
<td>18817</td>
<td>50677</td>
<td>1848</td>
</tr>
<tr>
<td>PokeMinn</td>
<td>-1848</td>
<td>42055</td>
<td>-14051</td>
<td>-15637</td>
<td>-62389</td>
<td>-58985</td>
<td>-12719</td>
<td>-18817</td>
<td></td>
<td>34299</td>
<td>-12010</td>
</tr>
<tr>
<td>Manitoba-2</td>
<td>-2471</td>
<td>-5016</td>
<td>-22116</td>
<td>-10791</td>
<td>-52325</td>
<td>-40256</td>
<td>-27040</td>
<td>-50677</td>
<td>-34299</td>
<td></td>
<td>-27221</td>
</tr>
</tbody>
</table>
No-Limit is what you see on TV - bets can be any size

This was our first time making a No-Limit bot

Took third place, using a CFR bot with abstracted betting
No-Limit is what you see on TV - bets can be any size
This was our first time making a No-Limit bot
Took third place, using a CFR bot with abstracted betting
We hope to do better next year! Lots of exciting work to be done here.

<table>
<thead>
<tr>
<th></th>
<th>BluffBot20</th>
<th>GS3</th>
<th>Hyperborean07</th>
<th>SlideRule</th>
<th>Gomel</th>
<th>Gomel-2</th>
<th>Milano</th>
<th>Manitoba</th>
<th>PokeMinn</th>
<th>Manitoba-2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>BluffBot20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS3</td>
<td>-267</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-3036</td>
</tr>
<tr>
<td>Hyperborean07</td>
<td>-380</td>
<td>-113</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SlideRule</td>
<td>-576</td>
<td>-503</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gomel</td>
<td>-2093</td>
<td>-3161</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gomel-2</td>
<td>-2885</td>
<td>-124</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milano</td>
<td>-3437</td>
<td>-1875</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manitoba</td>
<td>-475</td>
<td>-4204</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PokeMinn</td>
<td>-1848</td>
<td>42055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manitoba-2</td>
<td>-2471</td>
<td>-5016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-27221</td>
</tr>
</tbody>
</table>
First Man-Machine Poker Championship

- Beating human experts is a big milestone
Beating human experts is a big milestone
Tough to get statistical significance against humans
Beating human experts is a big milestone
Tough to get statistical significance against humans
So we played two at once with the same cards
Beating human experts is a big milestone
Tough to get statistical significance against humans
So we played two at once with the same cards
Four matches of 500 hands each
Beating human experts is a big milestone
Tough to get statistical significance against humans
So we played two at once with the same cards
Four matches of 500 hands each
Have to be ahead by 25 small bets to win a match
Phil Laak

- Background: Mechanical Engineer
- Started gambling in competitive backgammon
- Competes in the world’s biggest poker tournaments
Background: Computer consultant

Started out by playing...

Plays in $1000-$2000 Limit games
Background: Computer consultant
Started out by playing... Magic: The Gathering
Plays in $1000-$2000 Limit games
Background: Computer consultant

Started out by playing... Magic: The Gathering

Plays in $1000-$2000 Limit games

(This is a lot of money!)
We had 10 different bots to use:

- Several Counterfactual Regret Minimization approximate Nash equilibria
- Flavours of Restricted Nash Response counter-strategies
We had 10 different bots to use:
- Several Counterfactual Regret Minimization approximate Nash equilibria
- Flavours of Restricted Nash Response counter-strategies

We wanted a baseline to compare future bots against

Bot used: Mr. Pink, our finest abstraction CFR approximate Nash equilibrium
On Stage: Ali Eslami

Hotel: Phil Laak
Day 1, Session 1

On Stage: Ali Eslami

- Ali: $395
- Phil: -$465
- Polaris ends ahead by $70
- Result: Tie

Hotel: Phil Laak
Score so far: 1 Tie

The careful choice (Mr. Pink) did OK, so let's try something crazy!
Score so far: 1 Tie
The careful choice (Mr. Pink) did OK, so let's try something crazy!
Bot used: Mr. Orange / Crazy 8s
It's a CFR approximate Nash equilibrium in a broken game that encourages aggression
Hotel: Ali Eslami

On Stage: Phil Laak
Hotel: Ali Eslami

- Ali: -$2495
- Phil: $1570
- Polaris ends ahead by $925
- Result: Win

On Stage: Phil Laak
Score so far: 1 Win, 1 Tie

Which of our 10 bots to use this time?
Score so far: 1 Win, 1 Tie

Which of our 10 bots to use this time?

We pulled an all nighter and ran importance sampling on the last 1000 hands

Predicted the best 3 bots to use against each player
Score so far: 1 Win, 1 Tie
Which of our 10 bots to use this time?
We pulled an all nighter and ran importance sampling on the last 1000 hands
Predicted the best 3 bots to use against each player
Used a coach that chose between these 3 during the match
Day 2, Session 1

Hotel: Ali Eslami

On Stage: Phil Laak
Day 2, Session 1

Hotel: Ali Eslami

- Ali: -$635
- Phil: $1455
- Polaris ends behind by $820
- Result: Loss

On Stage: Phil Laak
Score so far: 1 Win, 1 Tie, 1 Loss

Decided to play it safe and go for a tie
Day 2, Session 2

- Score so far: 1 Win, 1 Tie, 1 Loss
- Decided to play it safe and go for a tie
- Bot used: Mr. Pink, the approximate Nash equilibrium from the first match
Onstage: Ali Eslami

Hotel: Phil Laak
Onstage: Ali Eslami

- Ali: $460
- Phil: $110
- Polaris ends behind by $570
- Result: Loss

Hotel: Phil Laak
Man-Machine Match Conclusions

- Very close game — we lost by 0.01 small bets/game, less than the tie margin

Ali: “This was not a win for us...I played the best heads-up poker I've ever played...we just barely won”

Post-game analysis (DIVAT) suggests that we outplayed them

We'd like to do another match next year

There's lots of exciting work to do here, too!
Man-Machine Match Conclusions

- Very close game — we lost by 0.01 small bets/game, less than the tie margin
- Ali: “This was not a win for us...I played the best heads-up poker I’ve ever played...we just barely won”
Man-Machine Match Conclusions

- Very close game — we lost by 0.01 small bets/game, less than the tie margin
- Ali: “This was not a win for us...I played the best heads-up poker I’ve ever played...we just barely won”
- Post-game analysis (DIVAT) suggests that we outplayed them
Very close game — we lost by 0.01 small bets/game, less than the tie margin

Ali: “This was not a win for us...I played the best heads-up poker I’ve ever played...we just barely won”

Post-game analysis (DIVAT) suggests that we outplayed them

We’d like to do another match next year
Very close game — we lost by 0.01 small bets/game, less than the tie margin

Ali: “This was not a win for us...I played the best heads-up poker I've ever played...we just barely won”

Post-game analysis (DIVAT) suggests that we outplayed them

We’d like to do another match next year

There’s lots of exciting work to do here, too!