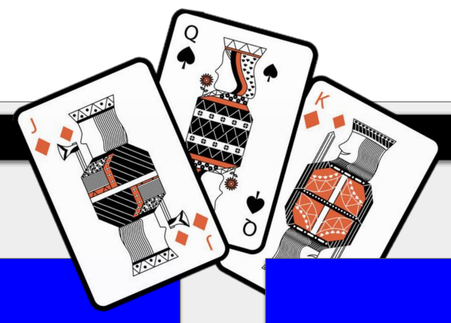
Training Poisoning in Imperfect Information Games

Natania Wolansky[1], Jisha Jacob[1], Guy Aridor[2], Iddo Drori[1]
[1] Department of Computer Science, Columbia University
[2] Department of Economics, Columbia University

COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science



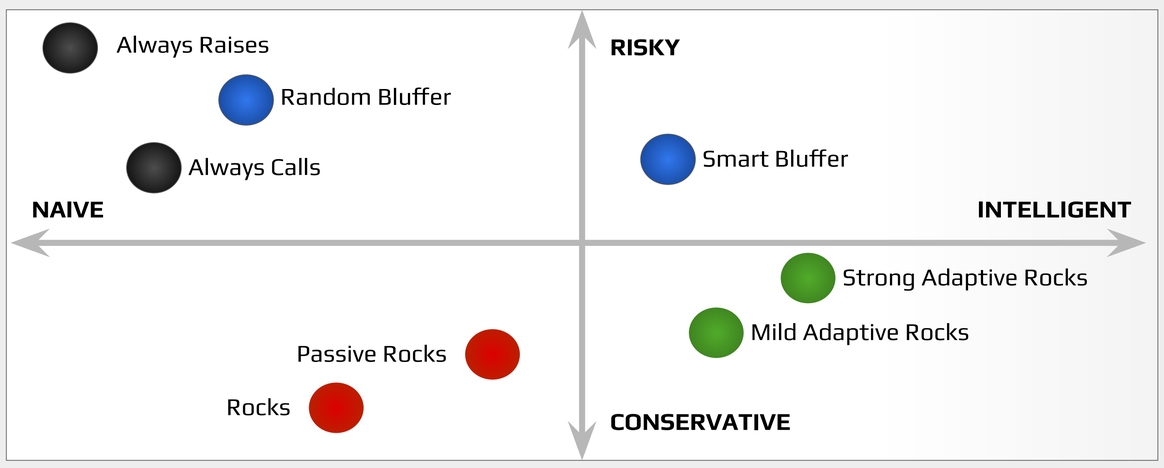
Summary & Goals

This work explores how simple strategies in the game of Leduc Hold'em can be used to beat a sophisticated pokerAI, DeepStack. We first analyze, under unbiased training,how significantly DeepStack outperforms most traditional poker-playing strategy profiles employed by humans.

We then consider the ability of an opponent to bias the training phase such that DeepStack is optimized to play against a particular strategy profile. Finally, by allowing for this biasing, we show that DeepStack can be defeated by a subset of strategy profiles if the player can change their strategy post-training. While DeepStack achieves nearly super-human performance, we conclude that DeepStack is susceptible to training poisoning.

Leduc Hold'em Poker

Semi-Rational & Irrational Players

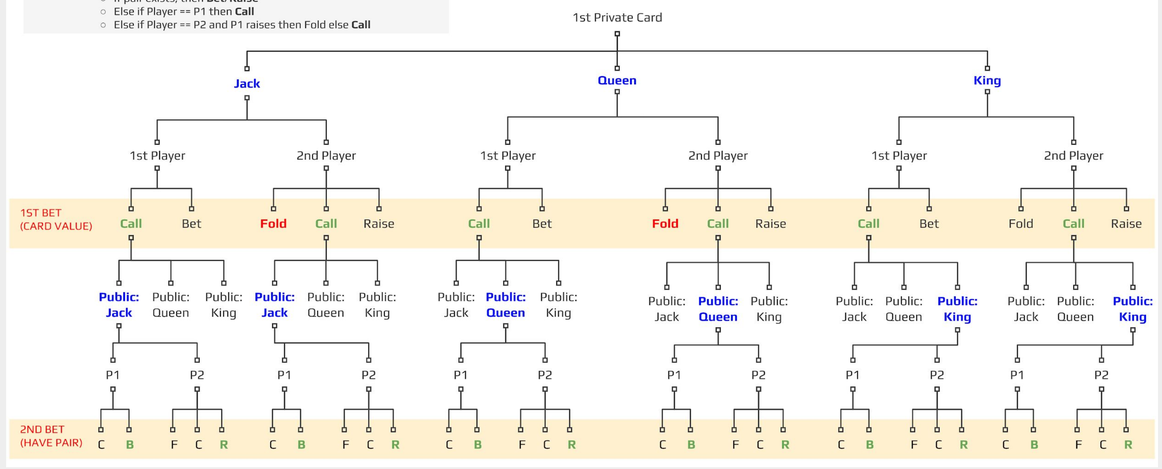


Rocks Player Game Tree

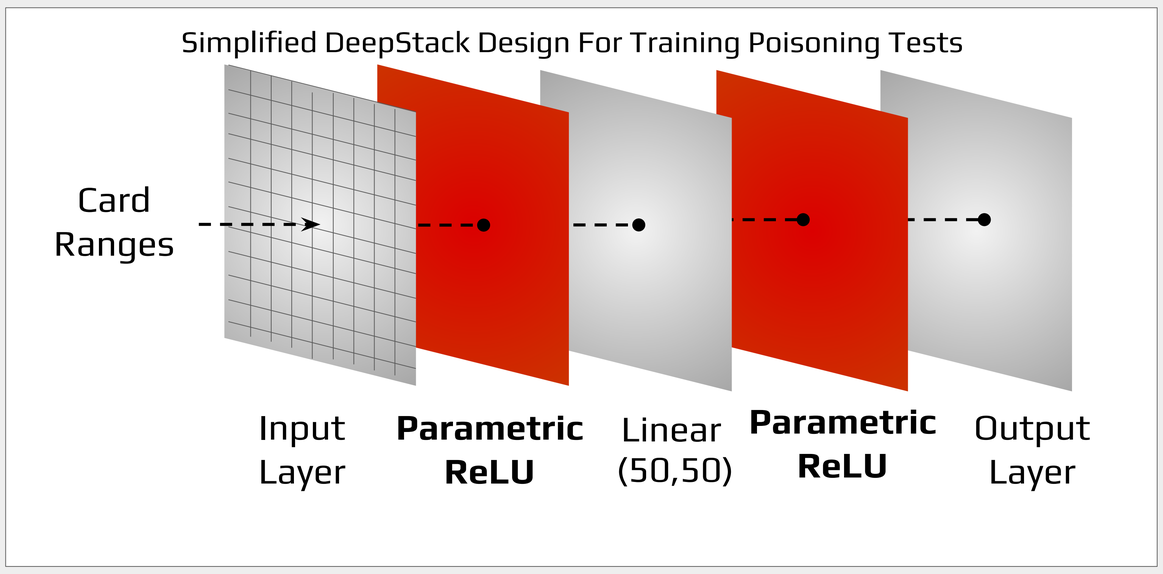
Risk Profile 1 (Rocks): Has a Pair

Rules:

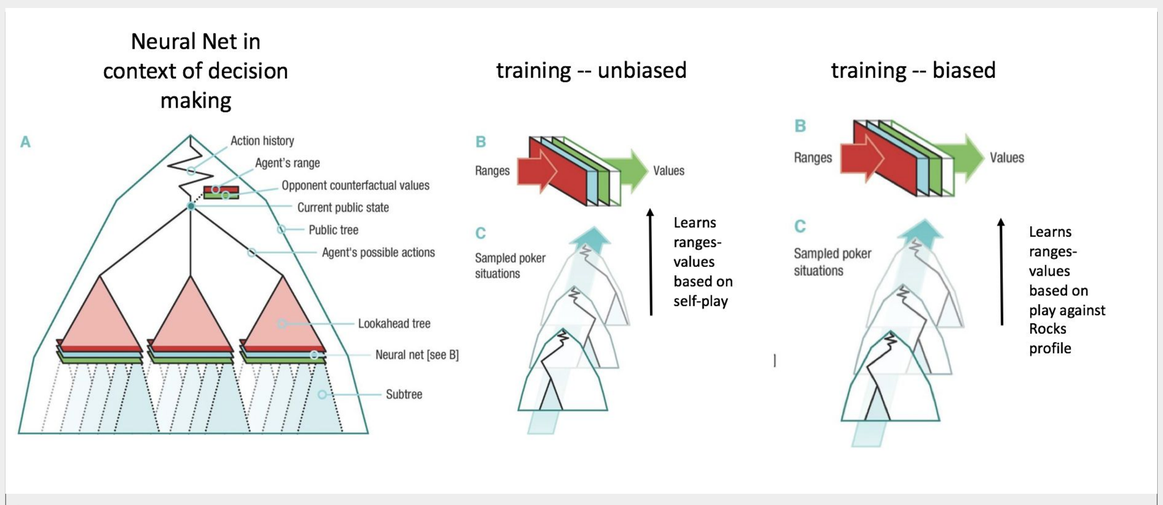
- 1st Bet:
  - If Player == P1, then **Call** (if P2 raises then use P2 logic below)
  - If Player == P2:
    - If Private card == K then **Call**
    - Else:
      - If P1 raises, **Fold**
      - If P1 calls, **Call**
- 2nd Bet:
  - If pair exists, then **Bet/Raise**
  - Else if Player == P1 then **Call**
  - Else if Player == P2 and P1 raises then Fold else **Call**



DeepStack Architecture



Training Poisoning in DeepStack



Source: Moravčík, et. al (2017)

Results

	Player	Unbiased	Biased
1	Mild Adaptive Rocks	-11.4 ± 12	63.6 ± 28
2	Passive Rocks	-43.5 ± 20	-14.8 ± 38
3	Strong Adaptive Rocks	-3.7 ± 9.4	53.9 ± 29
4	Rocks	-1.5 ± 5.8	4.7 ± 18
5	Random Bluffer	-53.6 ± 30	6 ± 34
6	Smart Bluffer	-35.8 ± 24	23.6 ± 32

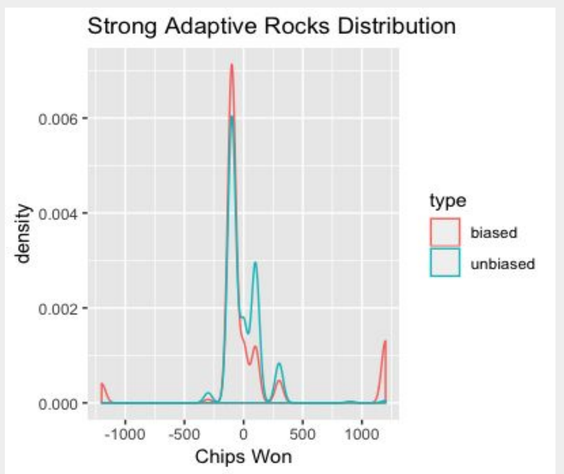
Table 2. Average Chips Per Game on Biased vs. Unbiased Training. The table reports means and 95% confidence intervals.

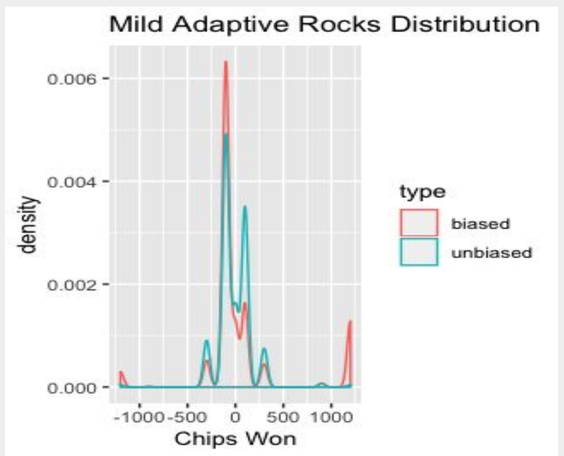
	Player	Unbiased	Biased
1	Mild Adaptive Rocks	192	457
2	Passive Rocks	328	620
3	Strong Adaptive Rocks	152	468
4	Rocks	93.9	296
5	Random Bluffer	492	544
6	Smart Bluffer	393	517

Table 3. Standard Deviation of Chips Won

	Player	Unbiased	Biased
1	Mild Adaptive Rocks	0.371	0.291
2	Passive Rocks	0.455	0.434
3	Strong Adaptive Rocks	0.327	0.253
4	Rocks	0.418	0.331
5	Random Bluffer	0.586	0.466
6	Smart Bluffer	0.51	0.417

Table 4. Average Win Rate (Fraction of Rounds Winning $>$ 0 Chips) - Biased vs Unbiased





Source code: github.com/rawls238/LeducTrainingPoisoning/