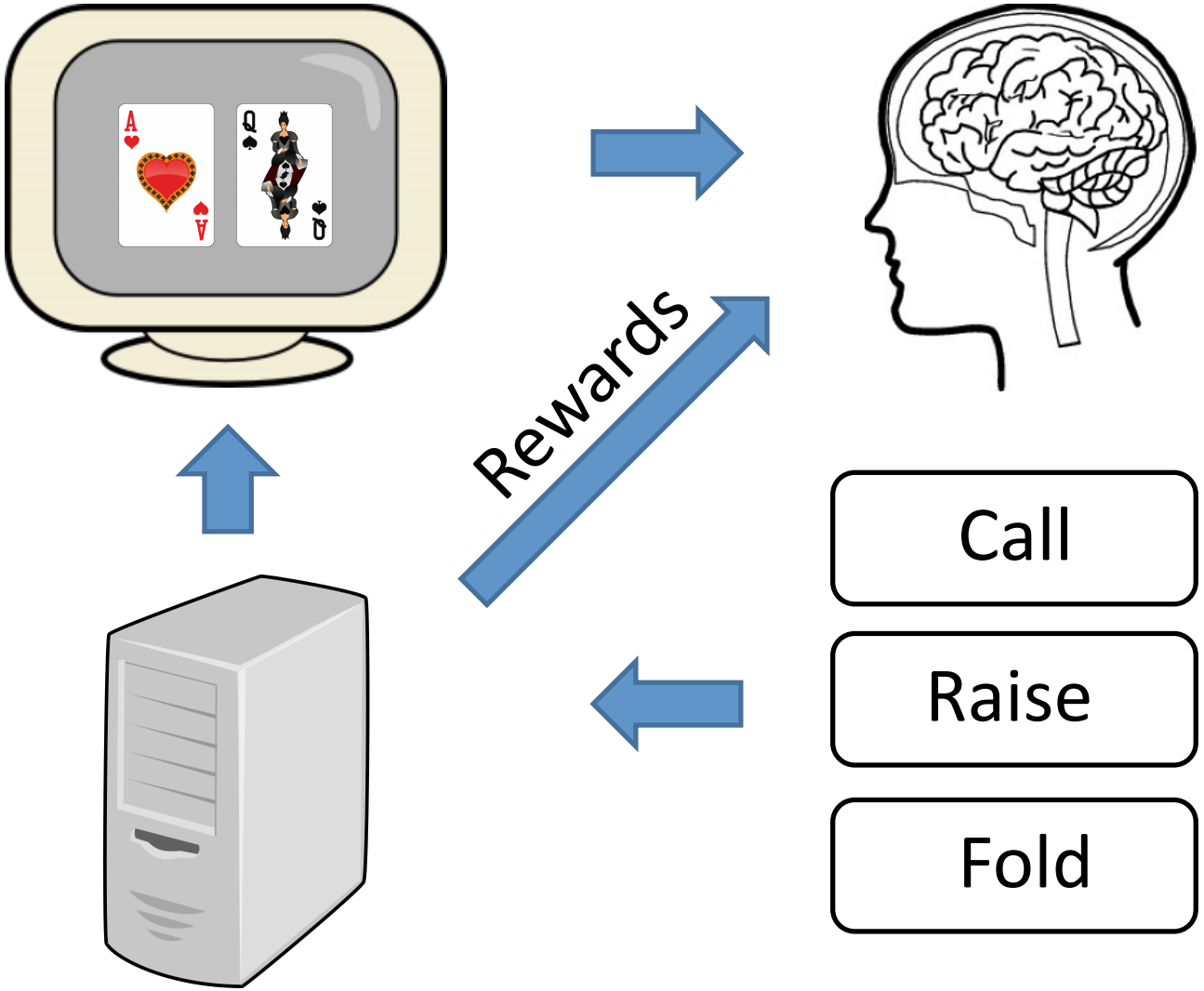
Poker-CNN: Learning Pattern in Poker Games Using a Convolutional Network

Nikolai Yakovenko – Twitter Cortex
Liangliang Cao – Yahoo Labs & Columbia University
Colin Raffel – Columbia University
James Fan – Columbia University
AAAI 2016

Poker is a Turn-Based Video Game



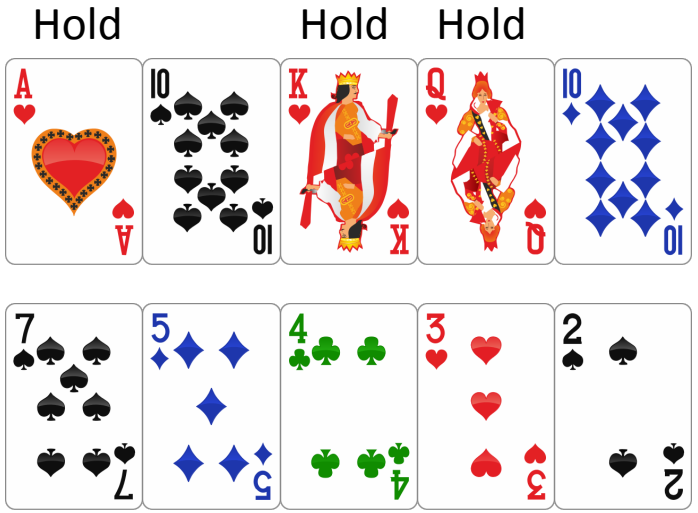
Unified Convet Poker Framework

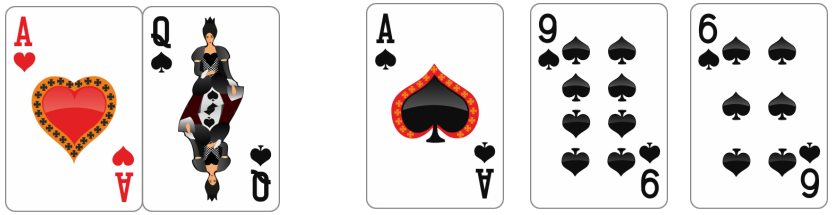
Single Draw Video Poker
98.5% payout

2-7 Lowball Triple Draw
Competitive against experts
(from self-play only)

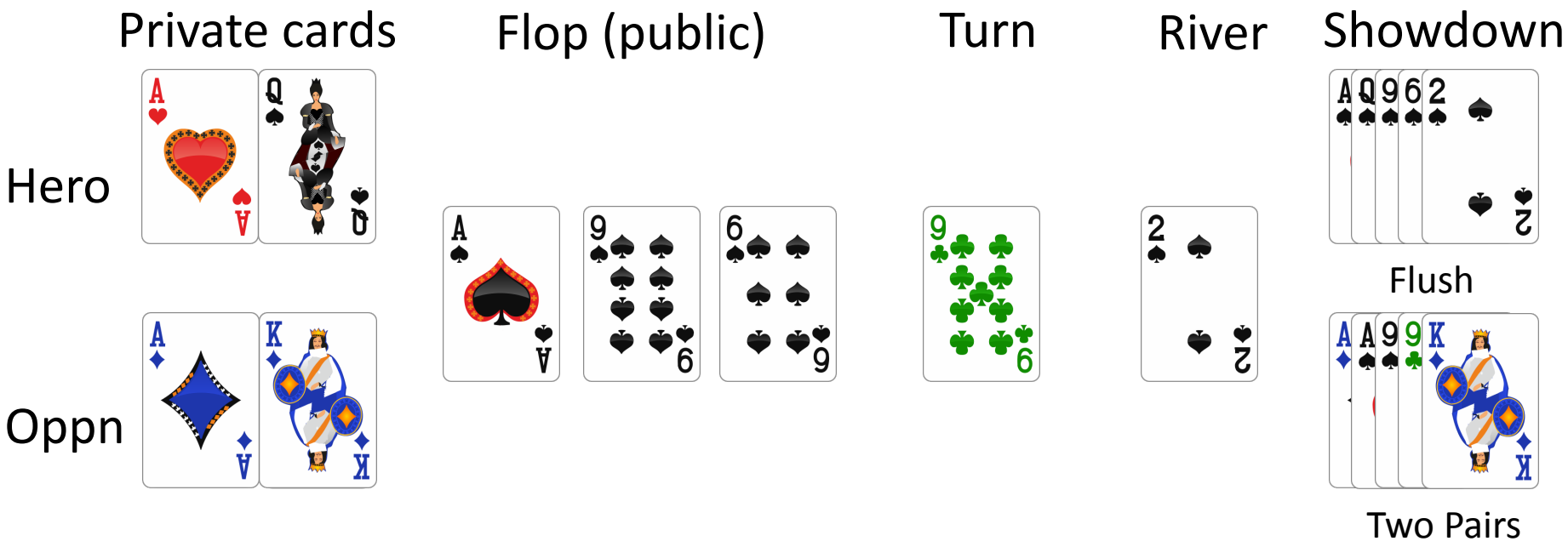
Limit Hold'em
Competitive against experts
(from self-play only)

No Limit Hold'em
Tied-5th 2016 ACPC
($29/hand behind the winner)





Texas Hold'em



CFR: Equilibrium Balancing

- Abstract Hold'em game to smaller state-space
- Cycle over ever game state.
- Converges to Nash equilibrium in simplified game.

- Winners of every Annual Computer Poker Competition (ACPC) since 2007.
  - Limit Hold'em: 1% of unexploitable
  - No Limit Hold'em: statistical tie vs pro players

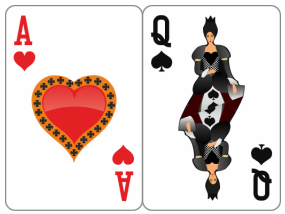
No Limit Hold'em Complexity

- 2016 ACPC 1st place winner Slumbot:
  - 9.7 x 10^11 information sets
  - Each state encoded as single byte
  - 2 months to train
  - $7,000 on AWS
- Tartanian (CMU)
  - Statistical tie vs best poker pros
  - 2 TB online lookup table

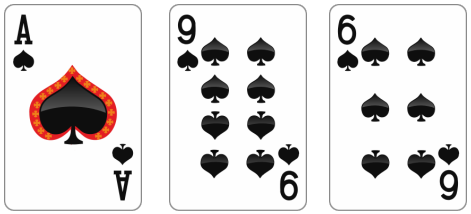
Can we train a strong poker player with a much smaller strategy?

Poker-CNN: Cards as 2D Tensors

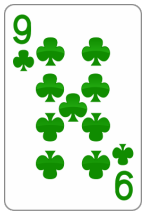
Private cards



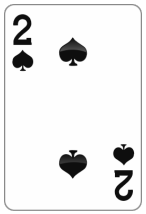
Flop (public)



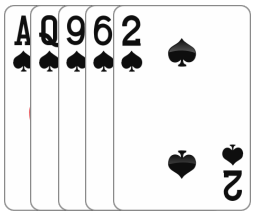
Turn



River



Showdown



Flush

```
[AhQs]

x23456789TJQKA
c.............
d.............
h............1
s..........1..
```

```
[AhQs]+[As9s6s]

x23456789TJQKA
c.............
d.............
h............1
s...1..1..1.1
```

Flush draw

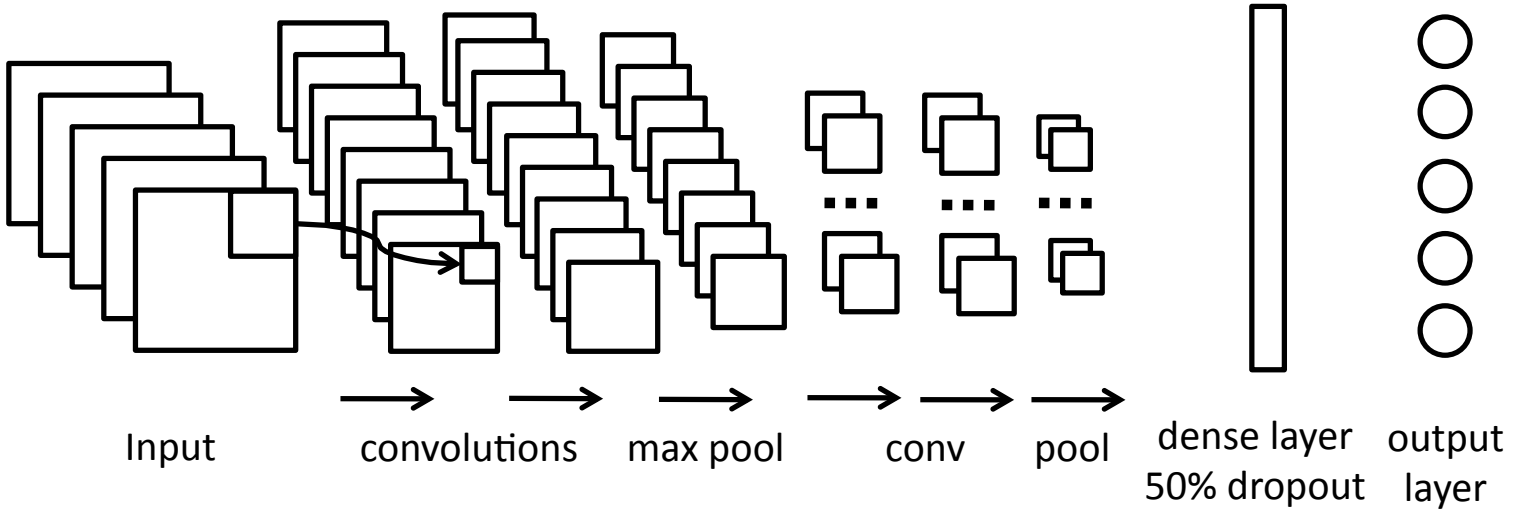
Pair (of Aces)

```
[AhQsAs9s6s9c2s]

x23456789TJQKA
c.......1.....
d.............
h............1
s1...1..1..1.1
```

Flush!

Convet: Predict Anything You Want



Inputs:

- Private cards
- Public cards
- Pot size
- Position
- Previous bets history

(31 x 17 x 17 3D tensor)

Predict action value:

- Bet, call, fold values
- Action probabilities
- Value by bet size

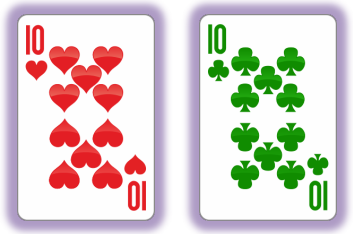
Surrogate tasks:

- Allin odds
- Opponent hand distribution

- single-trial $ win/loss

- no gradient for bets not made

- no Monte Carlo or tree search required

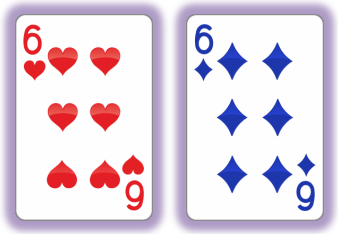


Big Blind

$100

Small Blind

$50



$20,000

+$265

$20,000

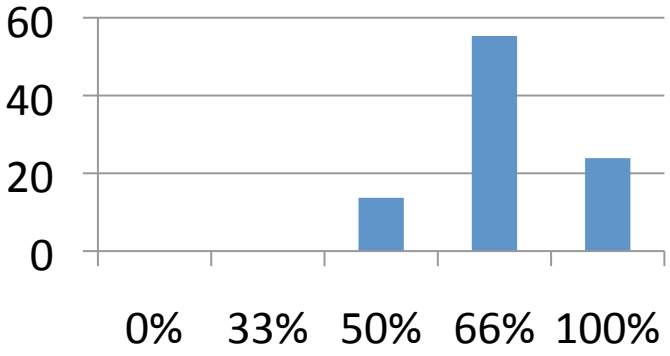
Raise	81.1%
Call	18.9%
Fold	0.0%

+$2616

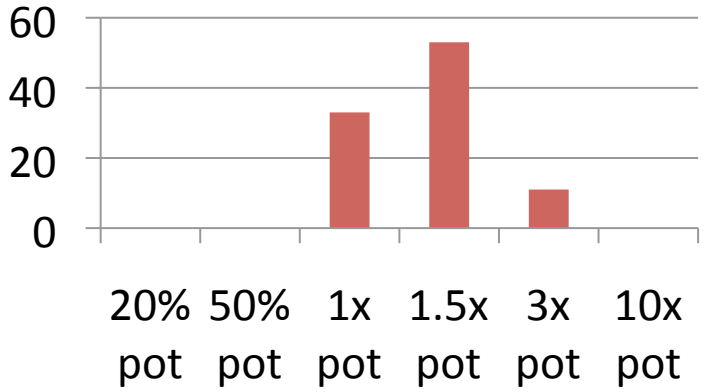
(Call)

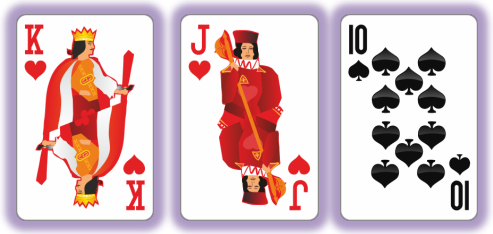
Raise	84.2%
Call	15.8%
Fold	0.0%

Odds vs Opponent

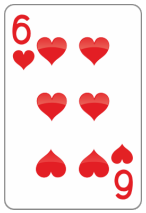


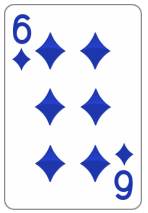
Bet Size





$5,430



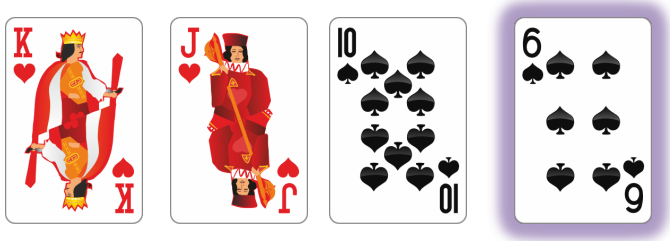


$17,285

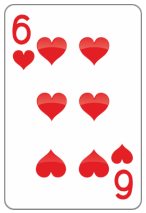
$17,285

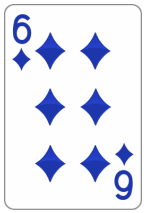
Bet	30.0%	(Check)	(Check)	Bet	25.9%
Check	70.0%			Check	74.1%

Value vs random	91.3%	Value vs random	52.9%
Value vs oppon	85.6%	Value vs oppon	32.6%



$5,430



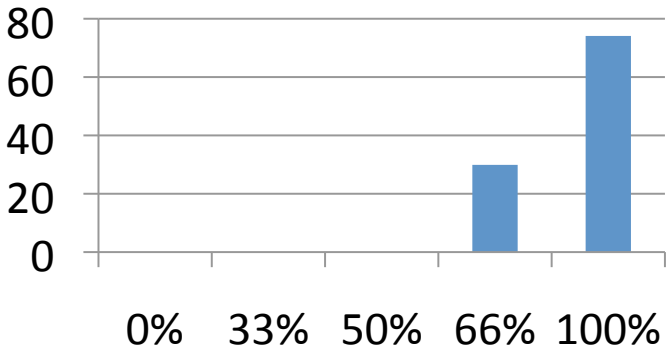


$17,285

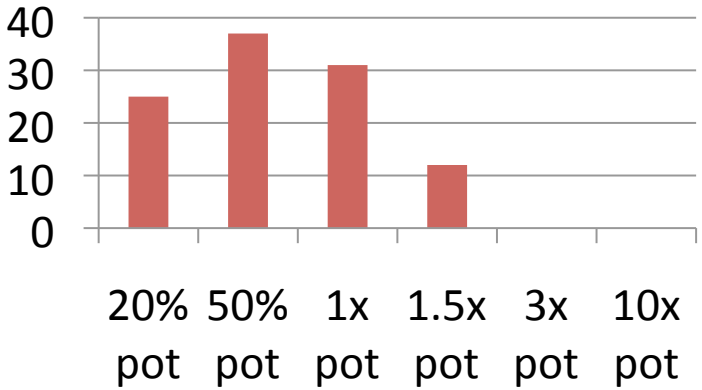
$17,285

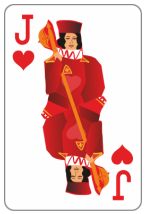
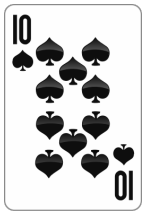
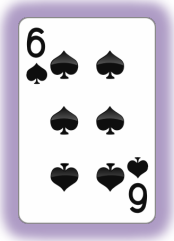
Bet	59.0%	(Check)	+$3,967	Bet	86.6%
Check	41.0%			Check	13.4%

Odds vs Opponent

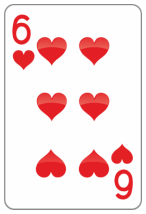
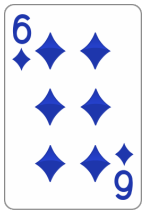


Bet Size




$9,397

$17,285

$13,318

+$3,967

Raise	26.6%	+$17,285 (allin)		Call	32.4%
Call	73.4%			Fold	67.6%
Fold	0.0%				
Value vs random	91.3%			Value vs random	84.7%
Value vs oppon	68.0%			Value vs oppon	30.6%

($13,000 allin call, to win $26,000)

33.3% odds = break-even

Takeaways

- Pattern matching pretty good, with enough data
  - Naïve network design
  - Training ≈ 4 million 2014 ACPC hands
  - No batch shuffle
  - No reinforcement learning
  - No LSTM memory units
- Struggles with rare cases
  - Under-weights outliers
  - Out of sample
- Struggles in big pots
  - Large importance for average results
  - Sparse data
- No attempt to avoid exploitability

Easy to Start, Easy to Modify

- Small 12-layer network in Theano
  - 2MB of model parameters
- Trains overnight on a single AWS GPU
- Easy to add surrogate tasks
  - For explanation
  - For better generalization
  - Small game-state abstractions for CFR?
- Relatively easy to train a new poker game

Future Work

- More games, more contexts
  - 3-player No Limit Hold'em
  - Pot-Limit Omaha
  - Tournament Hold'em
- Distill a CFR by learning internal parameters?
- Personal model against an opponent
- Hyper-parameter tuning…
  - 100,000 hands per experiment

References

- Paper on ArXiv http://arxiv.org/abs/1509.06731
- Code & models (admitted needs cleanup) at https://github.com/moscow25/deep_draw
- Annual Computer Poker Competition http://www.computerpokercompetition.org/
- Thanks to Eric Jackson of Slumbot for CFR code and advice.
- Thanks to poker pros Randy Ohel, Rep Porter & Neehar Banerji for help with human baselines.