Super-Human AI for Strategic Reasoning:

*Beating **Top** Pros in Heads-Up No-Limit Texas Hold'em*

Professor Tuomas Sandholm
Computer Science Department, CMU
Founder and CEO, Strategic Machine, Inc.

Noam Brown
Computer Science Department, CMU

STRATEGIC MACHINE, INC.

Imperfect-information games

Super-Human AI for Strategic Reasoning:

*Beating **Top** Pros in Heads-Up No-Limit Texas Hold'em*

Professor Tuomas Sandholm
Computer Science Department, CMU
Founder and CEO, Strategic Machine, Inc.

Noam Brown
Computer Science Department, CMU

STRATEGIC MACHINE, INC.

Imperfect-information games

Imperfect-information games

Imperfect-information games

Poker

Imperfect-information games

Poker

Security (Physical and Cyber)

AlphaGo

AlphaGo

AlphaGo techniques extend to all **perfect-information** games

Search in perfect-information games

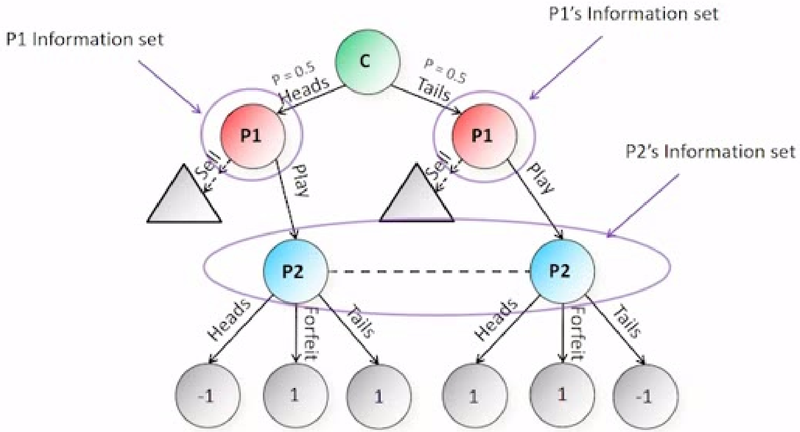
Search in perfect-information games

Sicilian Defense

Queen's Gambit

- An optimal response to the Queen's Gambit does not depend on the Sicilian Defense
- This is **not true** in imperfect-information games

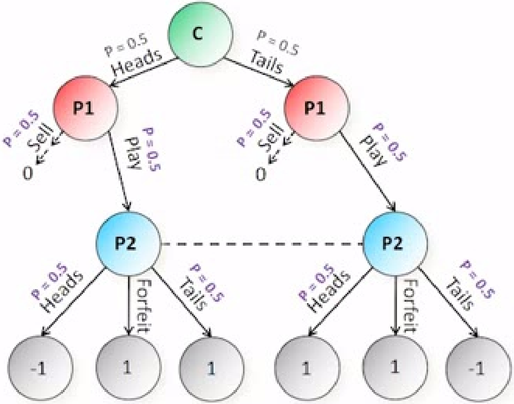
Imperfect-information games: Coin Toss



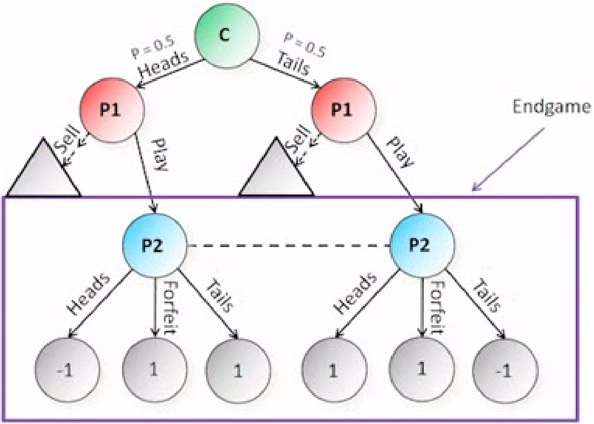
Nash equilibrium

Nash Equilibrium: a profile of strategies in which no player can improve by deviating (beliefs derived from strategies using Bayes rule)

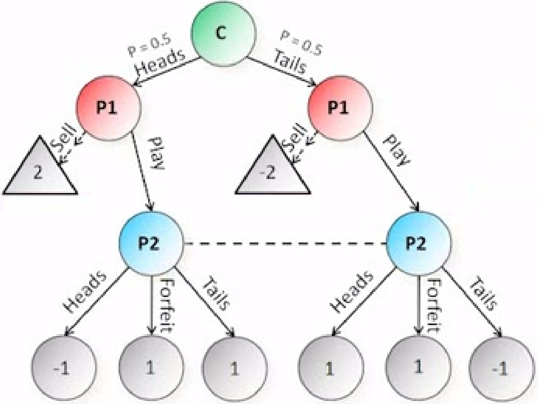
ϵ-Nash Equilibrium: No player can improve by more than ϵ



Imperfect-information games: Coin Toss



Imperfect-information games: Coin Toss



Tackling imperfect-info games

- Domain-independent techniques
- Techniques for complete-info games don't apply
- Challenges
  - Uncertainty about what others and chance will do
  - Unknown state => interpreting signals

Most real-world "games" are like this

- Negotiation
- Business strategy
- Strategic pricing
- Areas of finance
- Next-generation (cyber)security (zero-day vulnerabilities, jamming [DeBruhl et al.], OS)
- Political campaigns (e.g., media spending)
- Military (e.g., allocating troops, spending on space vs ocean, tactical)
- Auctions
- Steering evolution and biological adaptation, medical treatment planning [Sandholm 2012, AAAI-15 SMT Blue Skies]
- …

STRATEGIC MACHINE, INC.

Poker

- Recognized challenge problem in game theory and AI
  - [Nash 1950]
  - [Kuhn 1950]
  - [Waterman 1970]
  - [Caro 1984]
  - [Pfeffer & Koller 1995]
  - [Billings *et al.* 1998]
  - [Schaeffer *et al.* 1999]
  - [Shi & Littman 2001]
  - [Billings *et al.* 2003]
- Tremendous progress in the last 12 years
  - Rhode Island Hold'em solved ($\sim 10^9$ decisions) [Gilpin & Sandholm 2005]
  - Annual Computer Poker Competition started in 2016
  - Limit Texas Hold'em essentially solved ($\sim 10^{13}$ decisions) [Bowling *et al.* 2015]

Heads-up no-limit Texas hold’em

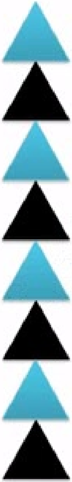
- Has become the main ***benchmark and challenge problem*** in AI for imperfect-information games

- 10^{161} situations

- Mostly played on the Internet
  - Also in World Series of Poker, NBC Heads-Up Championship, etc.
  - Featured in *Casino Royale* and *Rounders*

- No prior AI has been able to beat top humans

Texas hold'em



Chance deals 2 cards to each player

Round of betting

Chance deals 3 shared cards

Round of betting

Chance deals 1 shared card

Round of betting

Chance deals 1 shared card

Round of betting

Brains vs AI Rematch

- Libratus (= our AI) against four of the ***best*** heads-up no-limit Texas Hold'em specialist pros

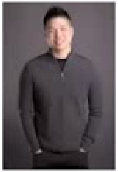
- 120,000 hands over 20 days in January 2017
- $200,000 divided among the pros based on performance

Conservative experiment design to favor humans

- Large number of hands
- Humans got to choose:
  - #days, break days, times of day, breaks between sessions—even dynamically
  - Two tabling
  - 4-color deck
  - Hot keys, adjustable dynamically
  - Specific hi-res monitors, their own mice
  - Twitch chat on vs off
  - Play in public vs private within each pair
- 200 big blinds deep
- No use of timing tells
- Action history displayed
- Hand histories given to both sides every evening, including hands opponent folded
- Humans allowed to:
  - Use computers and any programs to analyze
  - Collaborate and coordinate actions (except within each hand)
  - Get outside help (e.g., Doug Polk)
- Humans allowed to think as long as they want
- Mis-click hands canceled
- Ginseng ☺

Brains vs AI Rematch

- Libratus (= our AI) against four of the ***best*** heads-up no-limit Texas Hold'em specialist pros

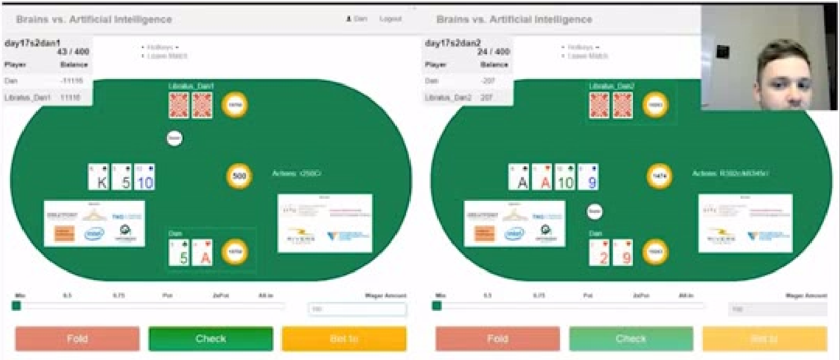


- 120,000 hands over 20 days in January 2017
- $200,000 divided among the pros based on performance

Conservative experiment design to favor humans

- Large number of hands
- Humans got to choose:
  - #days, break days, times of day, breaks between sessions—even dynamically
  - Two tabling
  - 4-color deck
  - Hot keys, adjustable dynamically
  - Specific hi-res monitors, their own mice
  - Twitch chat on vs off
  - Play in public vs private within each pair
- 200 big blinds deep
- No use of timing tells
- Action history displayed
- Hand histories given to both sides every evening, including hands opponent folded
- Humans allowed to:
  - Use computers and any programs to analyze
  - Collaborate and coordinate actions (except within each hand)
  - Get outside help (e.g., Doug Polk)
- Humans allowed to think as long as they want
- Mis-click hands canceled
- Ginseng ☺

User interface



Final result

- Libratus beat the top humans in this game by a lot
  - 147 mbb/hand
  - Statistical significance 99.98%, i.e., 0.0002
  - Each human lost to Libratus

Libratus's cumulative lead day by day

Final result

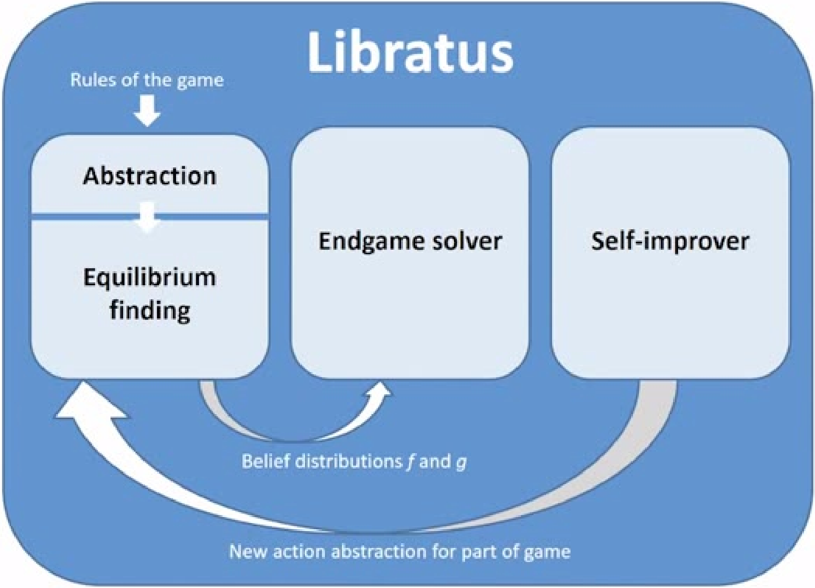
- Libratus beat the top humans in this game by a lot
  - 147 mbb/hand
  - Statistical significance 99.98%, i.e., 0.0002
  - Each human lost to Libratus

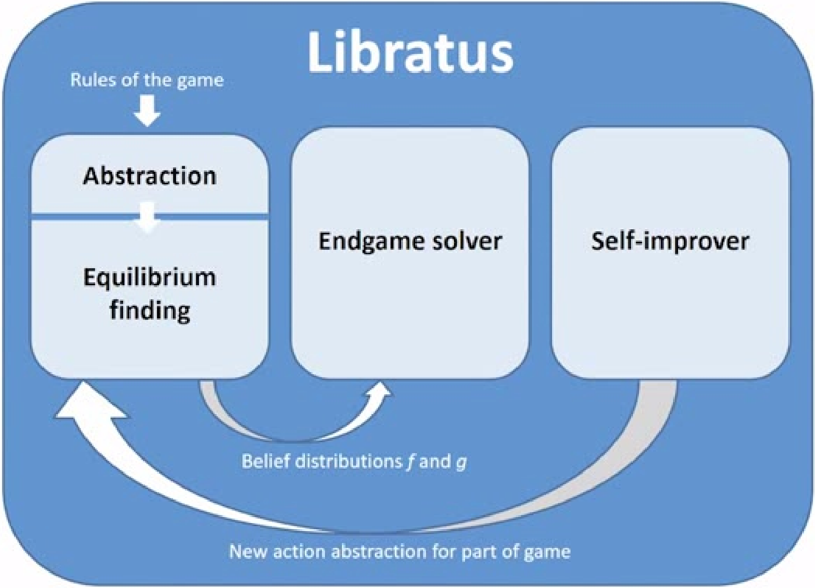
Libratus's cumulative lead day by day

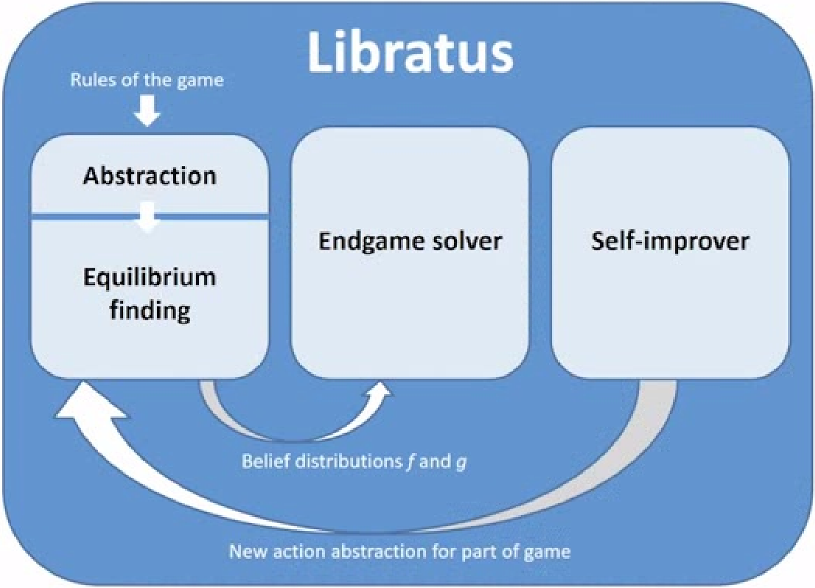
Lengpudashi vs humans event

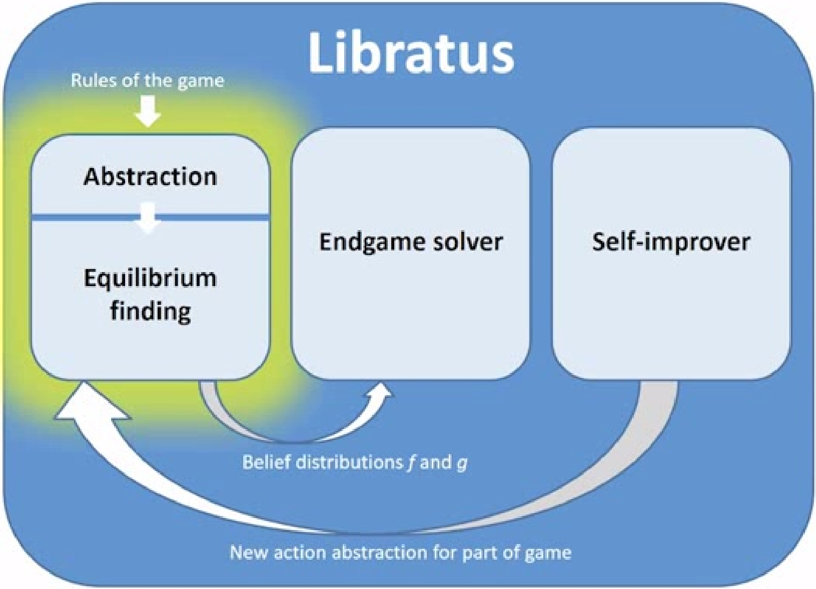
- 36,000 hands against 6 Chinese poker players
- Well-prepared opponents
  - WSOP bracelet winner
  - Expertize in computer science & ML
  - Studied Libratus's hand histories in advance
- 4.5 days: April 6-10, 2017
- Lengudashi won by 220 mbb/hand
  - Won each of the 9 sessions
  - Also beat each human individually

How do Libratus and Lengpudashi work?

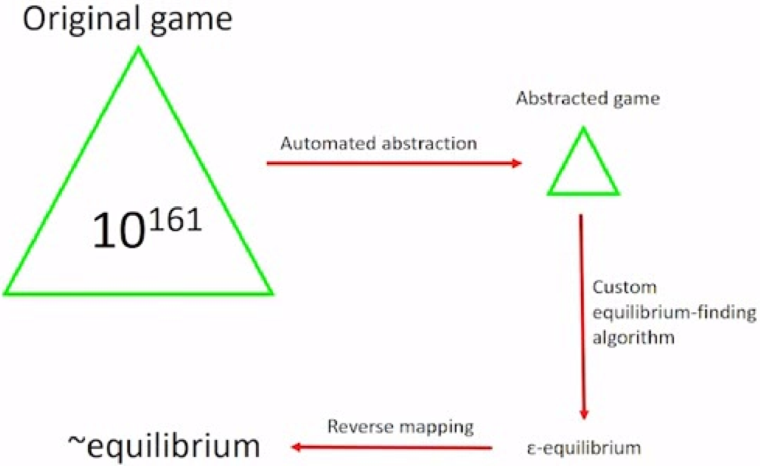

Libratus
Rules of the game
Abstraction
Equilibrium finding
Endgame solver
Self-improver
Belief distributions f and g
New action abstraction for part of game


Libratus
Rules of the game
Abstraction
Equilibrium finding
Endgame solver
Self-improver
Belief distributions *f* and *g*
New action abstraction for part of game


Libratus
Rules of the game
Abstraction
Equilibrium finding
Endgame solver
Self-improver
Belief distributions f and g
New action abstraction for part of game


Libratus
Rules of the game
Abstraction
Equilibrium finding
Endgame solver
Self-improver
Belief distributions f and g
New action abstraction for part of game

Abstraction [Gilpin & Sandholm EC-06, *J. of the ACM* 2007...]



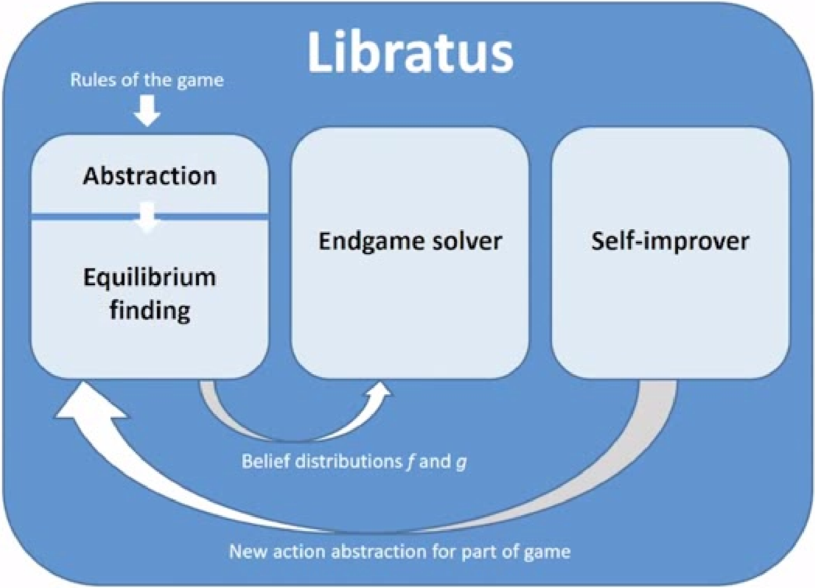
Foreshadowed by Shi & Littman 01, Billings et al. IJCAI-03

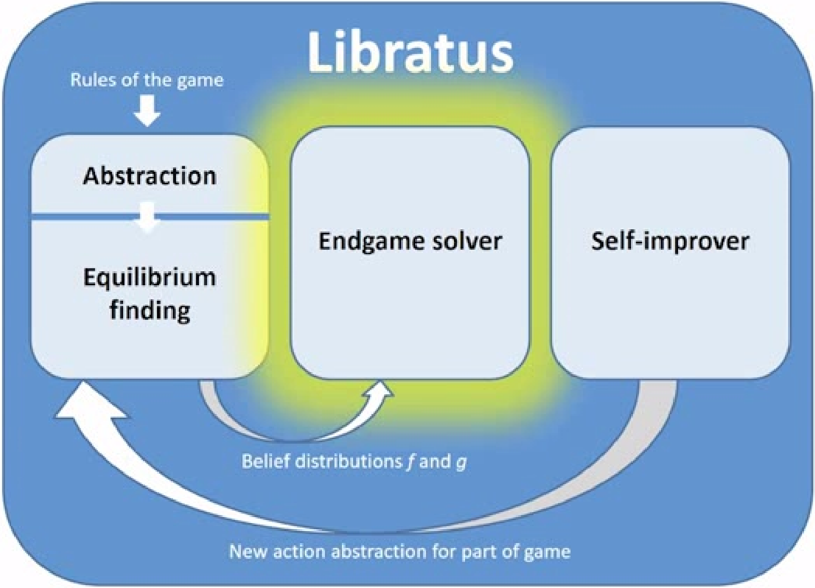
New equilibrium-finding algorithm

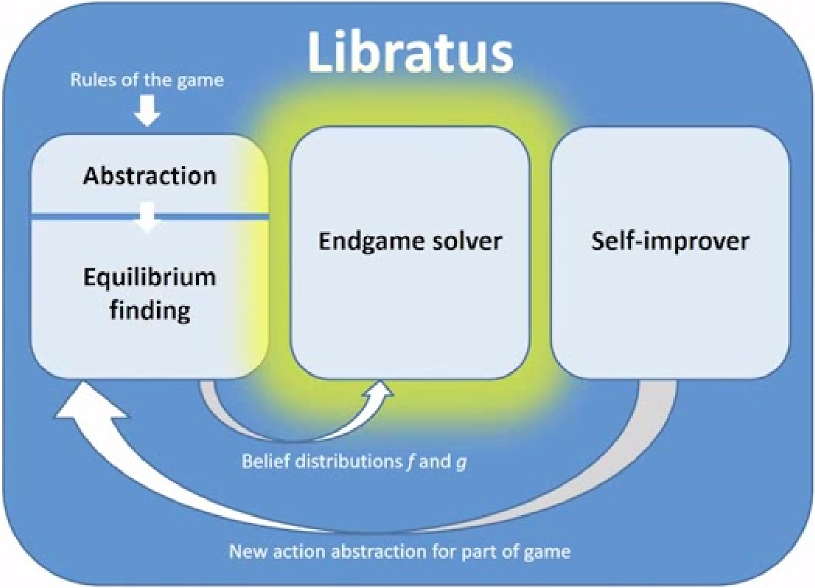
- Improvement on Monte-Carlo Counterfactual Regret Minimization [Lanctot *et al. NIPS*-09]
- Starts visiting less often paths where our own actions don't look promising (similar to Brown & Sandholm NIPS-15 paper and AAAI-17 workshop paper)
  => Speedup => can solve larger abstractions
- Also, the imperfect-recall abstraction, in effect, becomes finer grained
  => Better solution quality
- Distributed across 1 + 195 compute nodes

Card abstraction

- Same card abstraction algorithm that we used in Tartanian8
- But much finer abstraction
  - 1st and 2nd betting round: no abstraction
  - 3rd betting round: 55M card histories -> 2.5M buckets
  - 4th betting round: 2.4B card histories -> 1.25M buckets

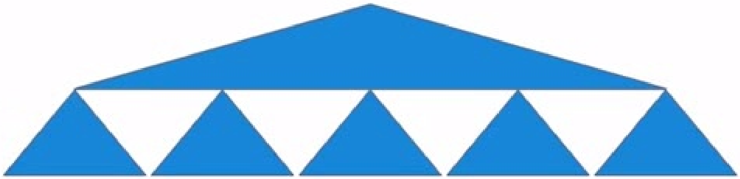

Libratus
Rules of the game
Abstraction
Equilibrium finding
Endgame solver
Self-improver
Belief distributions *f* and *g*
New action abstraction for part of game


Libratus
Rules of the game
Abstraction
Equilibrium finding
Endgame solver
Self-improver
Belief distributions *f* and *g*
New action abstraction for part of game


Libratus
Rules of the game
Abstraction
Equilibrium finding
Endgame solver
Self-improver
Belief distributions *f* and *g*
New action abstraction for part of game

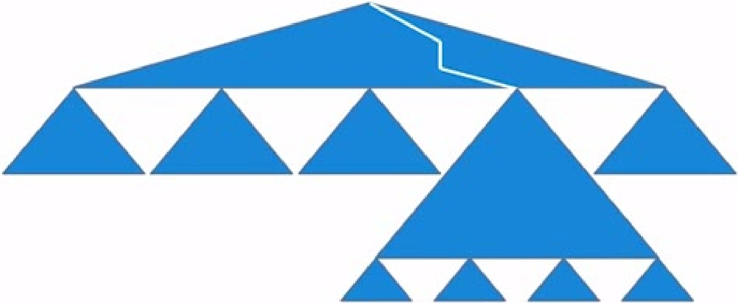
Endgame solving

[Burch *et al.* AAAI-14, Moravcik *et al.* AAAI-16, Brown & Sandholm AAAI-17]



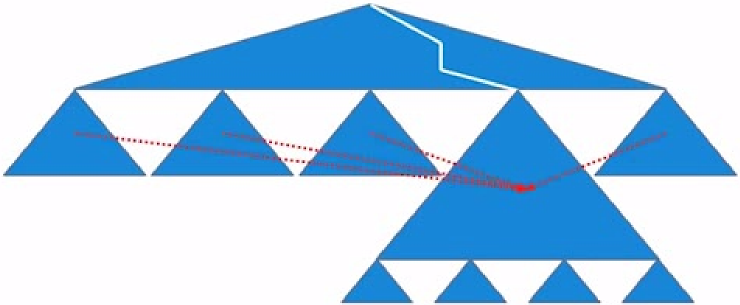
Endgame solving

[Burch *et al.* AAAI-14, Moravcik *et al.* AAAI-16, Brown & Sandholm AAAI-17]



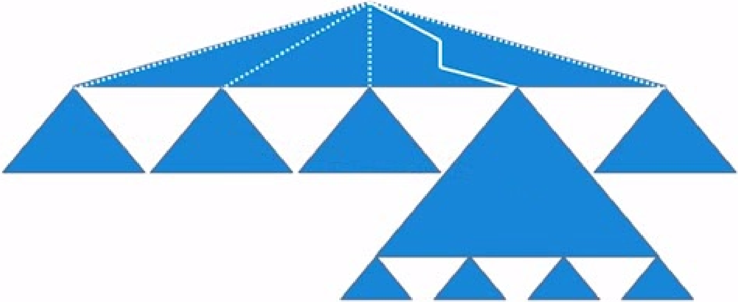
Endgame solving

[Burch *et al.* AAAI-14, Moravcik *et al.* AAAI-16, Brown & Sandholm AAAI-17]



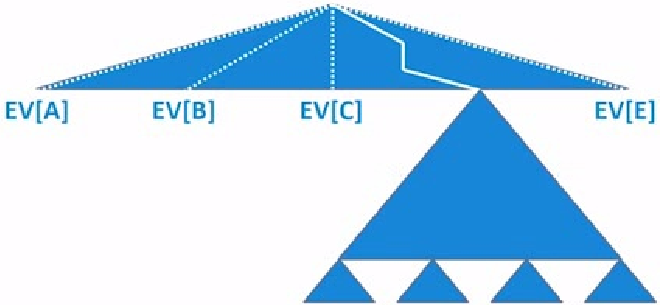
Endgame solving

[Burch *et al.* AAAI-14, Moravcik *et al.* AAAI-16, Brown & Sandholm AAAI-17]



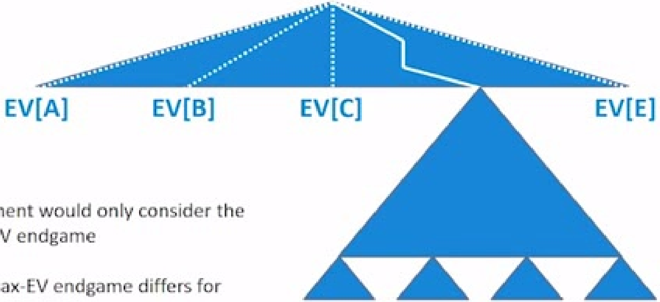
Endgame solving

[Burch *et al.* AAAI-14, Moravcik *et al.* AAAI-16, Brown & Sandholm AAAI-17]



Endgame solving

[Burch *et al.* AAAI-14, Moravcik *et al.* AAAI-16, Brown & Sandholm AAAI-17]

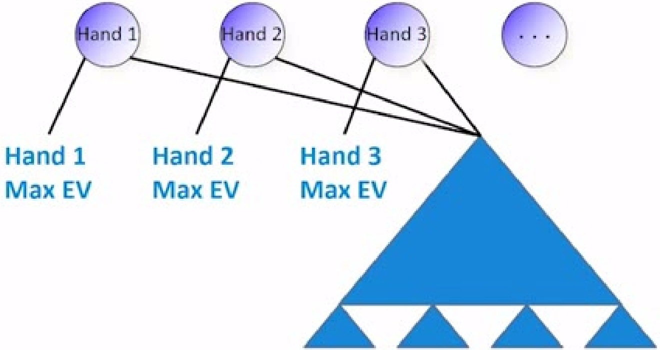


Opponent would only consider the max-EV endgame

The max-EV endgame differs for each poker hand

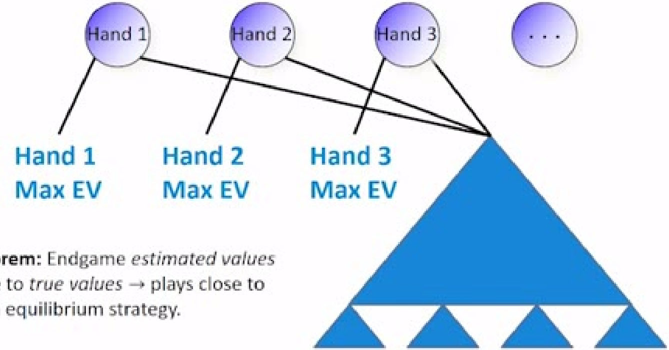
Endgame solving

[Burch *et al.* AAAI-14, Moravcik *et al.* AAAI-16, Brown & Sandholm AAAI-17]



Endgame solving

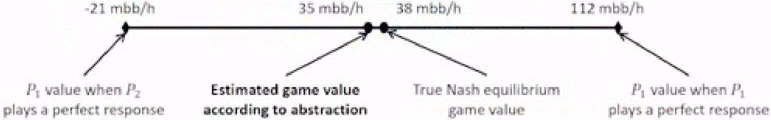
[Burch *et al.* AAAI-14, Moravcik *et al.* AAAI-16, Brown & Sandholm AAAI-17]



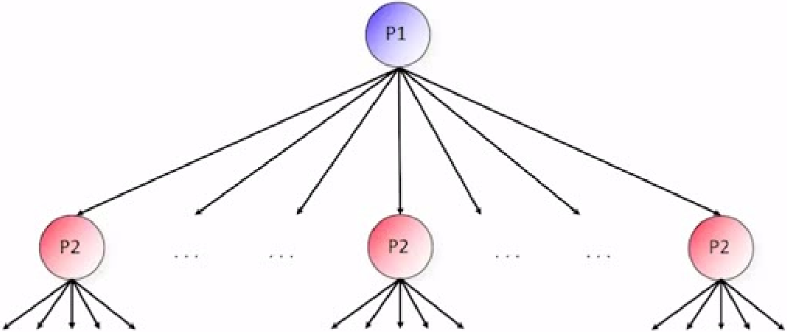
Theorem: Endgame *estimated values* close to *true values* → plays close to Nash equilibrium strategy.

How good are abstraction strategies?

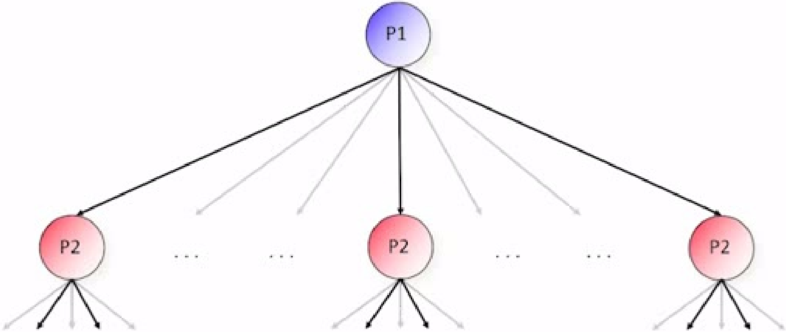
- Test game of Flop Texas Hold'em using an abstraction that is 0.02% of the full game size:



Action abstraction

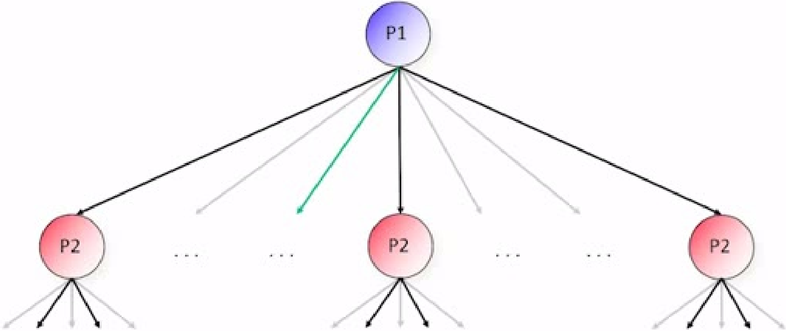


Action abstraction



[Gilpin *et al.* AAMAS-08]
[Hawkin *et al.* AAAI-11 AAAI-12]
[Brown & Sandholm AAAI-14]

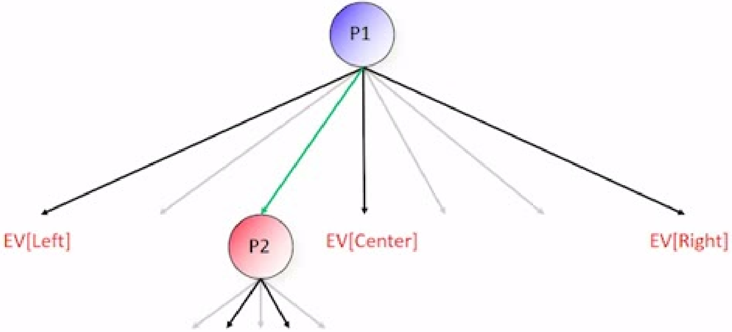
Action translation



[Gilpin *et al.* AAMAS-08]
[Schnizlein *et al.* IJCAI-09]
[Ganzfried & Sandholm IJCAI-13]

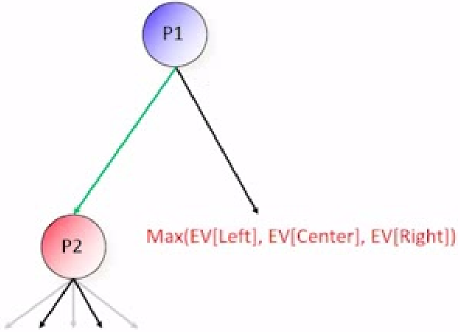
Nested endgame solving

[Brown & Sandholm AAAI-17]



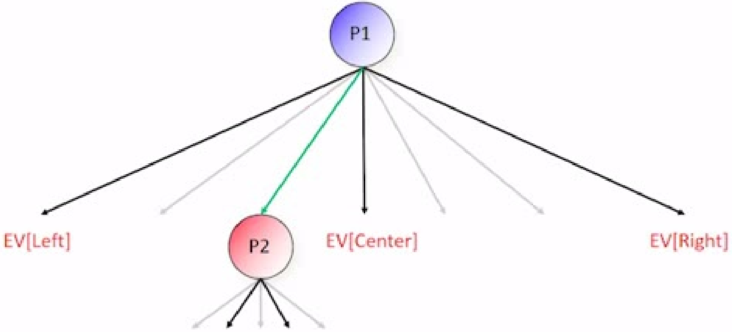
Nested endgame solving

[Brown & Sandholm AAAI-17]



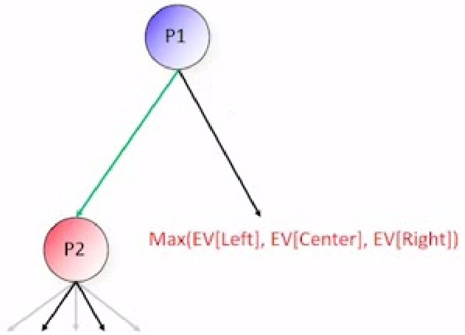
Nested endgame solving

[Brown & Sandholm AAAI-17]



Nested endgame solving

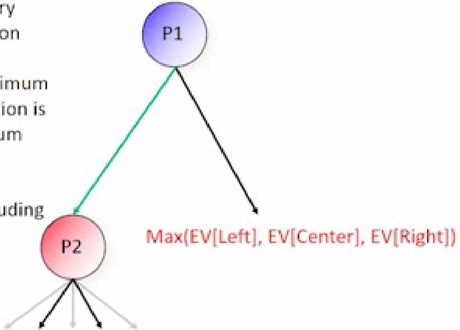
[Brown & Sandholm AAAI-17]



Nested endgame solving

[Brown & Sandholm AAAI-17]

- Can be repeated for every subsequent off-tree action
- Theoretically safe if maximum action EV in the abstraction is close to the *true* maximum action EV
- Can mitigate this by including *optimal* actions in the abstraction [Brown & Sandholm AAAI-14]

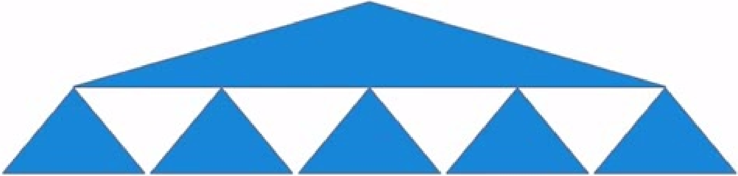


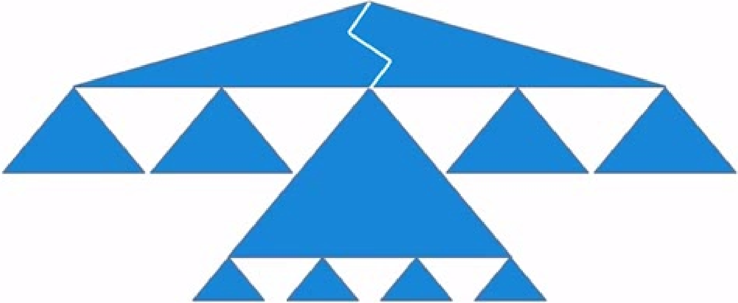
Medium-scale experiment on endgame solving within action abstraction

	Small Game Exploitability	Large Game Exploitability
Abstraction Strategy	91.3 mbb / hand	41.4 mbb / hand
Unsafe Endgame Solving	5.51 mbb / hand	397 mbb / hand
Safe Endgame Solving	**8.26 mbb / hand**	**5.50 mbb / hand**

Medium-scale experiments on nested endgame solving

	Exploitability
Randomized Pseudo-Harmonic Translation	1,465 mbb / hand
Nested Unsafe Endgame Solving	**148.3 mbb / hand**
Nested Safe Endgame Solving	**119.1 mbb / hand**





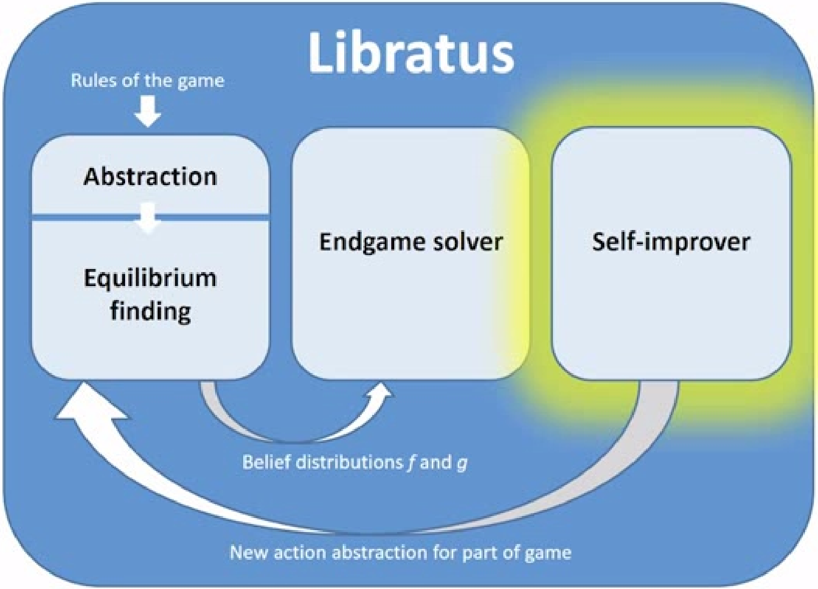
New ideas in endgame solver

- Safe endgame solving taking into account opponent's mistakes in the hand so far
- Nested endgame solving
- Endgame starts already on 3rd betting round
- No card abstraction in the endgame
- Noise added to action abstraction before solving

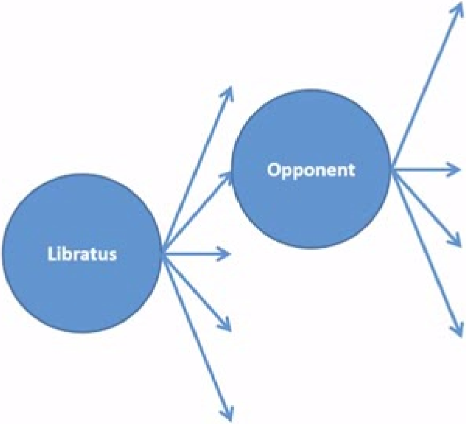
Daniel McAulay on Libratus's "balance" and use of "blockers"


The Fight For Humanity Rages On!
DANIEL MCAULAY
@678DMCA

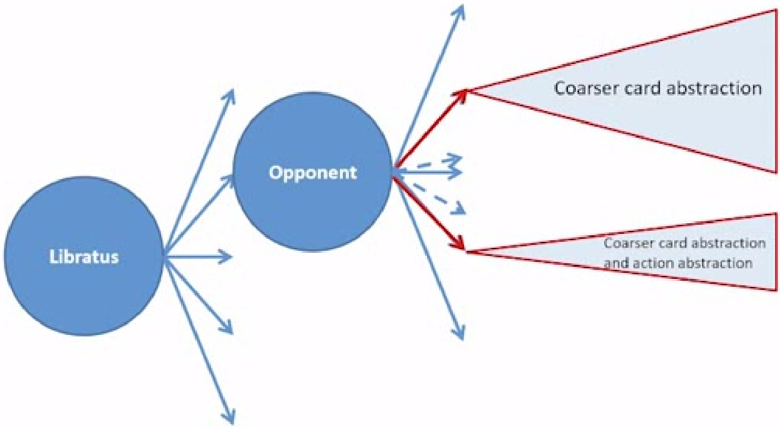
Daniel McAulay on Libratus's
"balance" and use of "blockers"


Libratus
Rules of the game
Abstraction
Equilibrium finding
Endgame solver
Self-improver
Belief distributions f and g
New action abstraction for part of game

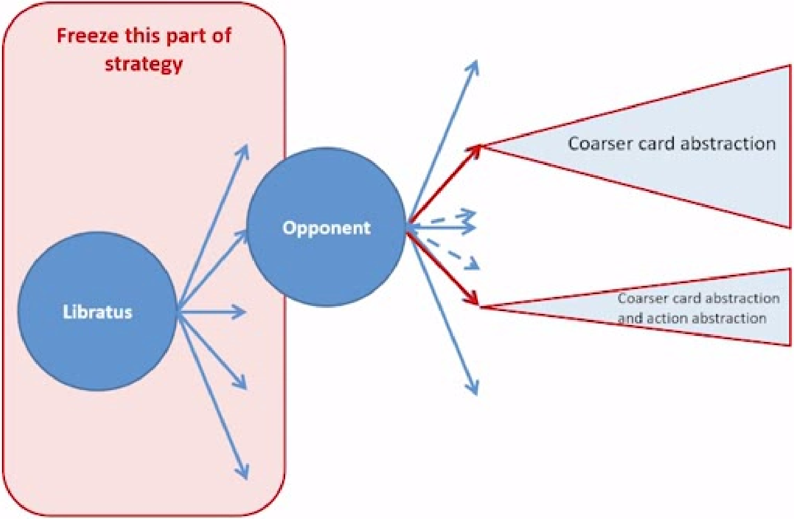
Filling holes in the betting tree



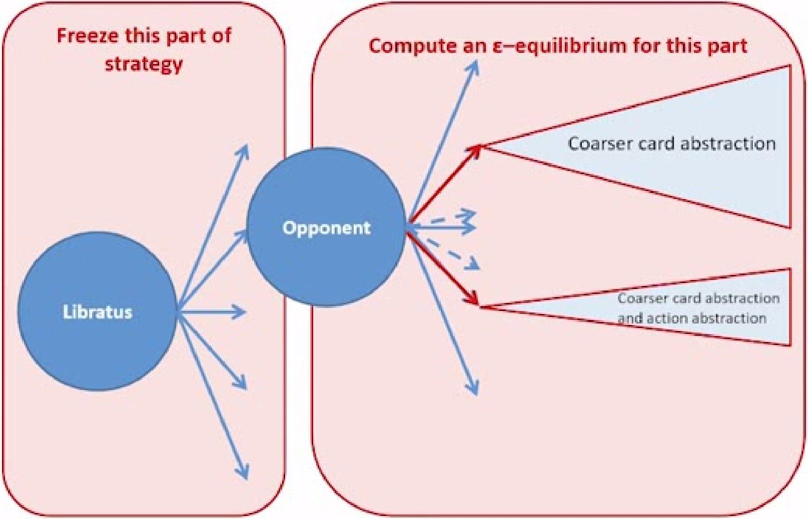
Filling holes in the betting tree



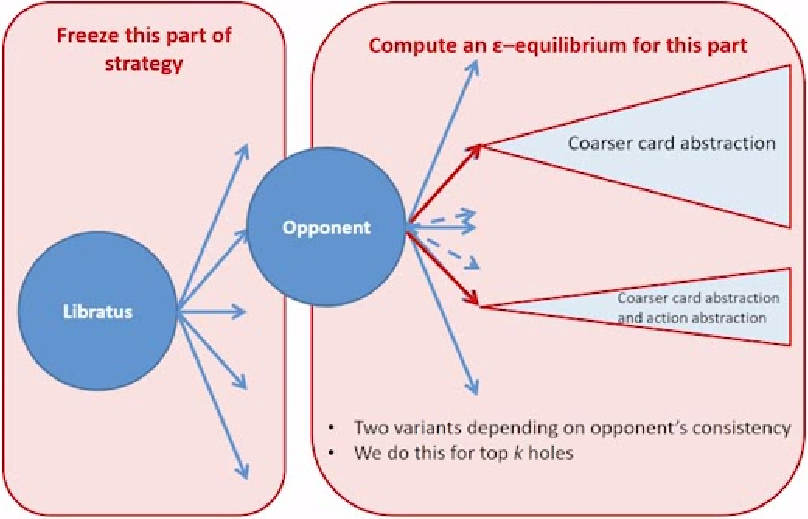
Filling holes in the betting tree



Filling holes in the betting tree



Filling holes in the betting tree



- Two variants depending on opponent's consistency
- We do this for top *k* holes

Jason Les and Jimmy Chou on Libratus's holes

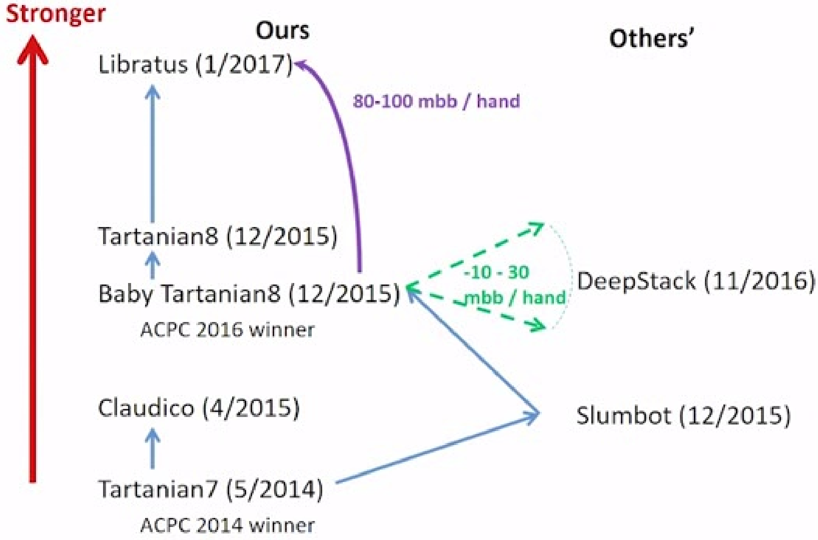

The Fight For Humanity Rages On!
JASON LES
@HEYITSCHEET

Jason Les and Jimmy Chou
on Libratus's holes


The Fight For Humanity Rages On!
JIMMY CHOU

Jason Les and Jimmy Chou on Libratus's holes

Head-to-head strength of recent AIs



Observations about Libratus's play

- Strengths:
  - Small bets & huge bets & huge all-ins
  - "Perfect balance"
  - Mixed strategy, not "range-based"
  - "Donk betting"
  - No card abstraction in parts that are played
  - Provably near-perfect endgame play
  - Different bet sizings used in endgames
- Weaknesses:
  - "No" opponent exploitation
- Safe (equilibrium) play = timid? Let's ask Dong & Dan


The Fight For Humanity Rages On!

Observations about Libratus's play

- Strengths:
  - Small bets & huge bets & huge all-ins
  - "Perfect balance"
  - Mixed strategy, not "range-based"
  - "Donk betting"
  - No card abstraction in parts that are played
  - Provably near-perfect endgame play
  - Different bet sizings used in endgames
- Weaknesses:
  - "No" opponent exploitation
- Safe (equilibrium) play = timid? Let's ask Dong & Dan

Some current & future research on this topic in my lab

- Practical lossy abstraction algorithms with bounds (also for modeling)
- New gradient-based equilibrium-finding algorithms [Kroer *et al.* EC-15, -17]
- Algorithms for equilibrium refinements [Kroer *et al.* IJCAI-17, Farina *et al.* 2017]
- Understanding exploration vs exploitation vs exploitability
- Additional applications
- …

Thank you!

Partners

TUOMAS SANDHOLM'S ELECTRONIC MARKETPLACES LABORATORY

Carnegie Mellon University
School of Computer Science

RIVERS CASINO

PITTSBURGH SUPERCOMPUTING CENTER

Sponsors

GREATPOINT
make no little plans

AVENUE 4 ANALYTICS

TNG TECHNOLOGY CONSULTING

Artificial Intelligence

intel

OPTIMIZED MARKETS

TNG TECHNOLOGY CONSULTING