First Man vs. Machine No-Limit Texas Hold 'em Competition

Sam Ganzfried

Assistant Professor, Computer Science, Florida International University, Miami FL

PhD, Computer Science Department, Carnegie Mellon University, 2015

sganzfri@cis.fiu.edu

Scope and applicability of game theory

- Strategic multiagent interactions occur in all fields
  - Economics and business: bidding in auctions, offers in negotiations
  - Political science/law: fair division of resources, e.g., divorce settlements
  - Biology/medicine: robust diabetes management (robustness against "adversarial" selection of parameters in MDP)
  - Computer science: theory, AI, PL, systems; national security (e.g., deploying officers to protect ports), cybersecurity (e.g., determining optimal thresholds against phishing attacks), internet phenomena (e.g., ad auctions)

Game theory background

	rock	paper	scissors
Rock	0,0	-1, 1	1, -1
Paper	1,-1	0, 0	-1,1
Scissors	-1,1	1,-1	0,0

- Players
- Actions (aka pure strategies)
- Strategy profile: e.g., (R,p)
- Utility function: e.g., $u_1(R,p) = -1$, $u_2(R,p) = 1$

Zero-sum game

	rock	paper	scissors
Rock	0,0	-1, 1	1, -1
Paper	1,-1	0, 0	-1,1
Scissors	-1,1	1,-1	0,0

- Sum of payoffs is zero at each strategy profile: e.g., $u_1(R,p) + u_2(R,p) = 0$
- Models purely adversarial settings

Mixed strategies

- Probability distributions over pure strategies
- E.g., R with prob. 0.6, P with prob. 0.3, S with prob. 0.1

Best response (aka nemesis)

- Any strategy that maximizes payoff against opponent's strategy
- If P2 plays (0.6, 0.3, 0.1) for r,p,s, then a best response for P1 is to play P with probability 1

Nash equilibrium

- Strategy profile where all players simultaneously play a best response
- Standard solution concept in game theory
  - Guaranteed to always exist in finite games [Nash 1950]
- In Rock-Paper-Scissors, the unique equilibrium is for both players to select each pure strategy with probability 1/3

Minimax Theorem

- Minimax theorem: For every two-player zero-sum game, there exists a value v* and a mixed strategy profile σ^* such that:
  a. P1 guarantees a payoff of at least v* in the worst case by playing σ^*_1
  b. P2 guarantees a payoff of at least -v* in the worst case by playing σ^*_2
- v* (= v_1) is the *value* of the game
- All equilibrium strategies for player i guarantee at least v_i in the worst case
- For RPS, v* = 0

Exploitability

- Exploitability of a strategy is difference between value of the game and performance against a best response
  - Every equilibrium has zero exploitability
- Always playing rock has exploitability 1
  - Best response is to play paper with probability 1

Nash equilibria in two-player zero-sum games

- Zero exploitability – “unbeatable”
- Exchangeable
  - If (a,b) and (c,d) are NE, then (a,d) and (c,b) are too
- Can be computed in polynomial time by a linear programming (LP) formulation

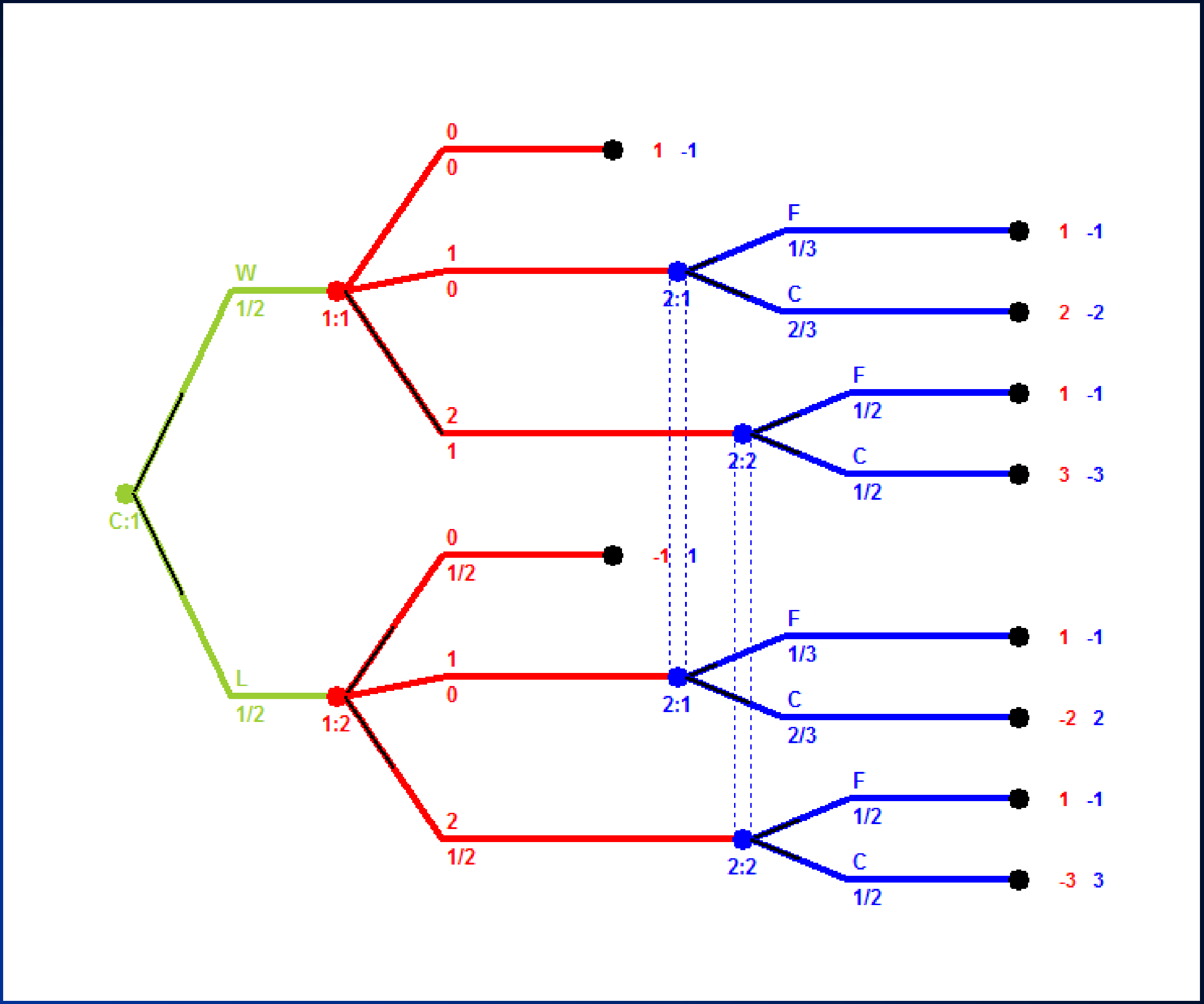
Nash equilibria in multiplayer and non-zero-sum games

- None of the two-player zero-sum results hold
- There can exist multiple equilibria, each with different payoffs to the players
- If one player follows one equilibrium while other players follow a different equilibrium, overall profile is not guaranteed to be an equilibrium
- If one player plays an equilibrium, he could do worse if the opponents deviate from that equilibrium
- Computing an equilibrium is PPAD-hard

Imperfect information

- In many important games, there is information that is private to only some agents and not available to other agents
  - In auctions, each bidder may know his own valuation and only know the distribution from which other agents' valuations are drawn
  - In poker, players may not know private cards held by other players

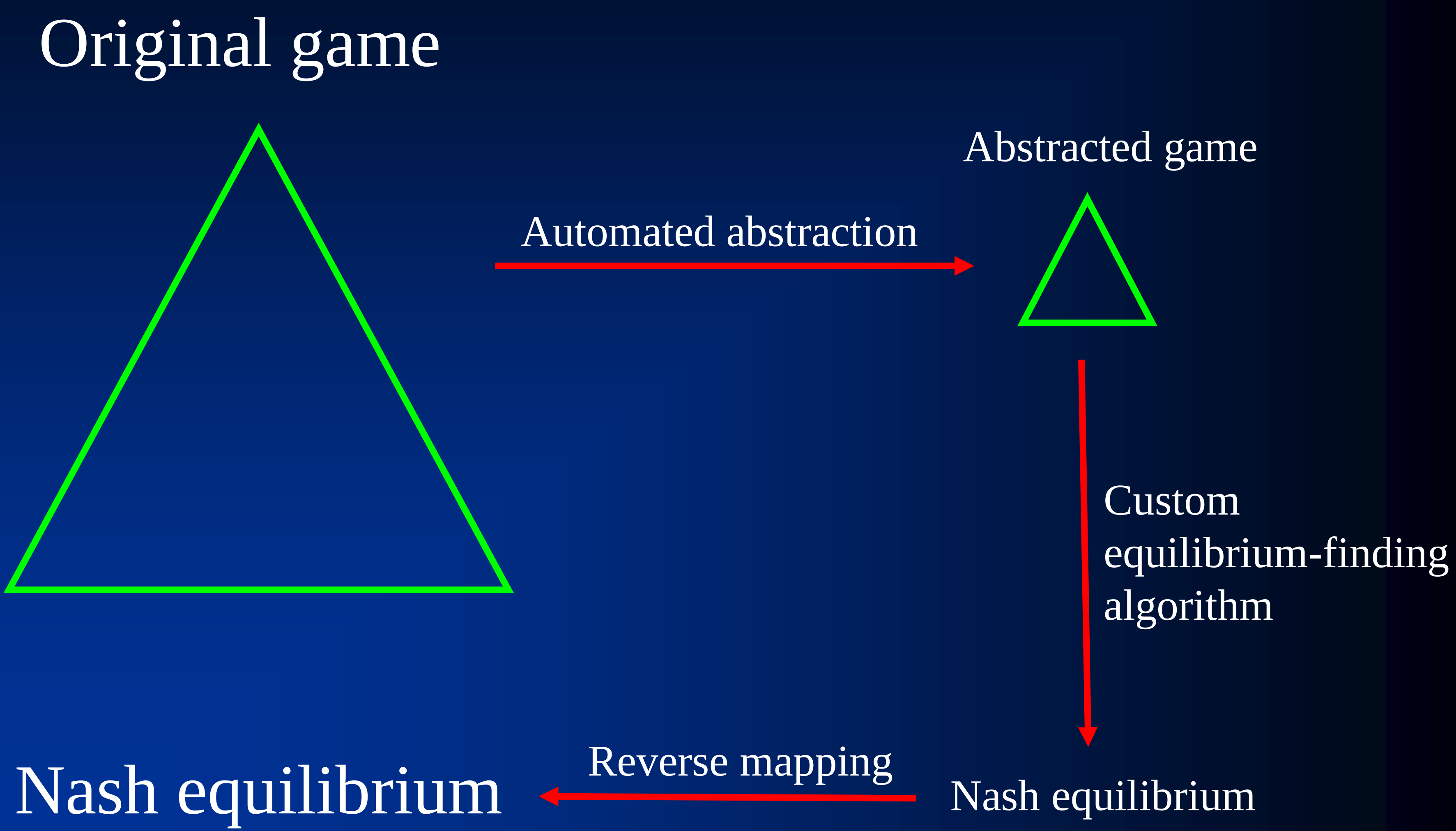
Extensive-form representation



Extensive-form games

- Two-player zero-sum EFGs can be solved in polynomial time by linear programming
  - Scales to games with up to 10^8 states
- Iterative algorithms (CFR and EGT) have been developed for computing an ε-equilibrium that scale to games with 10^{17} states
  - CFR also applies to multiplayer and general sum games, though no significant guarantees in those classes
  - (MC)CFR is self-play algorithm that samples actions down tree and updates regrets and average strategies stored at every information set

Standard paradigm for solving large imperfect-information games



Texas hold ‘em poker

- Huge game of imperfect information
  - Most studied imp-info game in AI community since 2006 due to AAAI computer poker competition
  - Most attention on 2-player variants (2-player zero-sum)
  - Multi-billion dollar industry (not “frivolous”)
- Limit Texas hold ‘em – fixed betting size
  - $\sim 10^{17}$ nodes in game tree
- No Limit Texas hold ‘em – unlimited bet size
  - $\sim 10^{165}$ nodes in game tree
  - Most active domain in last several years
  - Most popular variant for humans

No-limit Texas hold 'em poker

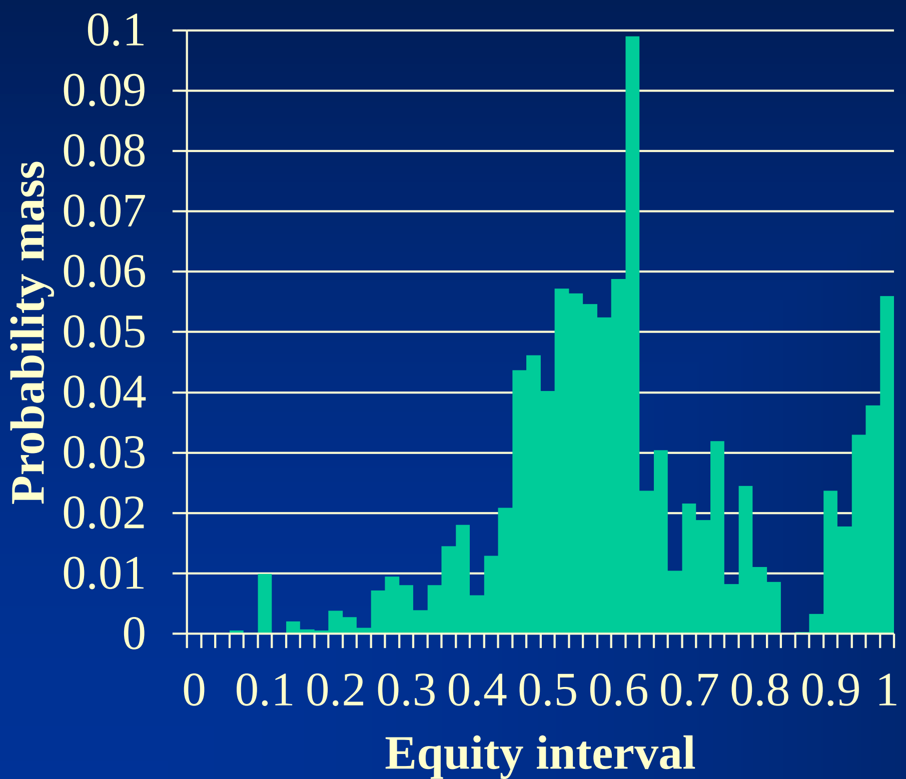
- Two players have stack and pay blinds (ante)
- Each player dealt two private cards
- Round of betting (preflop)
  - Players can fold, call, bet (any amount up to stack)
- Three public cards dealt (flop) and a second round of betting
- One more public card and round of betting (turn)
- Final card and round of betting (river)
- Showdown

Game abstraction

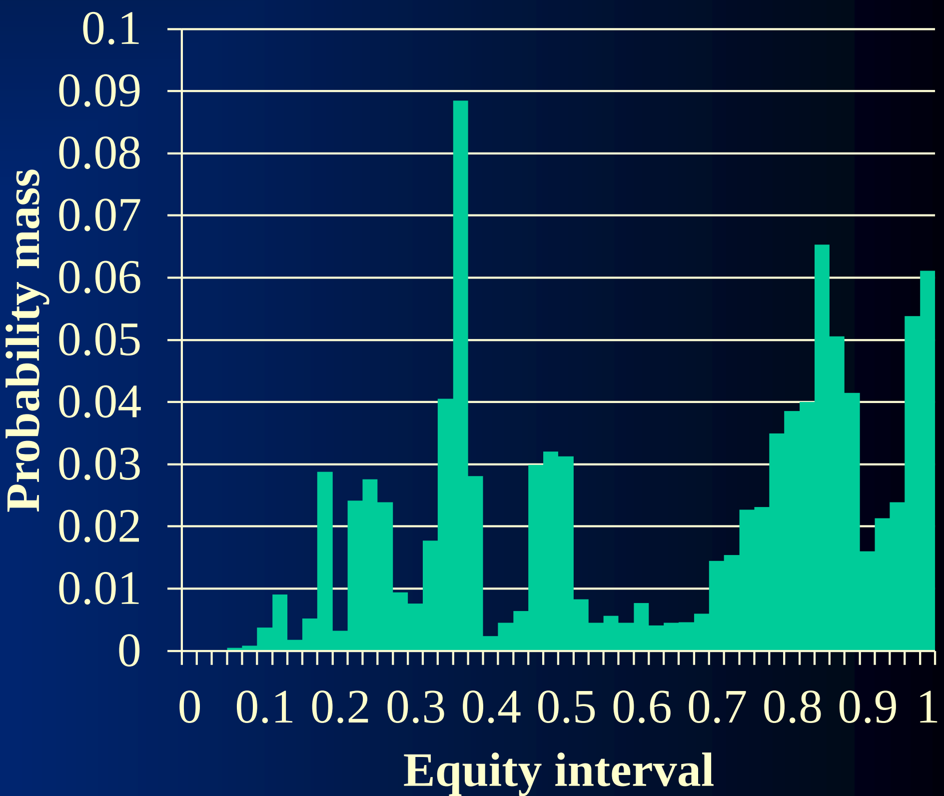
- Necessary for solving large games
  - 2-player no-limit Texas hold 'em has 10^{165} game states, while best solvers "only" scale to games with 10^{17} states
- Information abstraction: grouping information sets together
- Action abstraction: discretizing action space
  - E.g., limit bids to be multiples of \$10 or \$100

Information abstraction

Equity distribution for 6c6d. EHS: 0.634

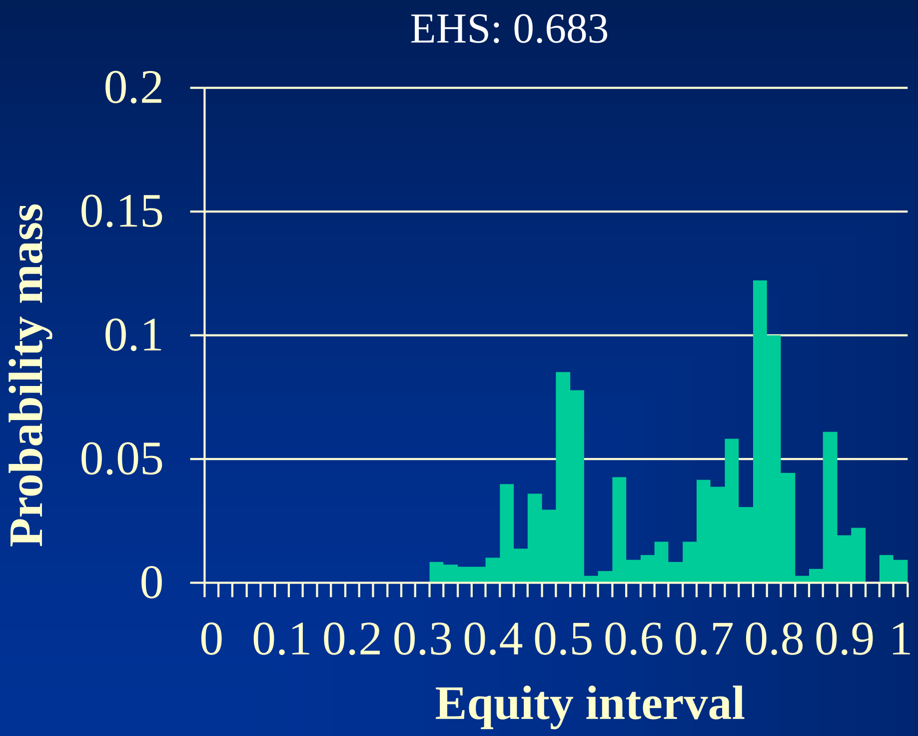


Equity distribution for KcQc. EHS: 0.633

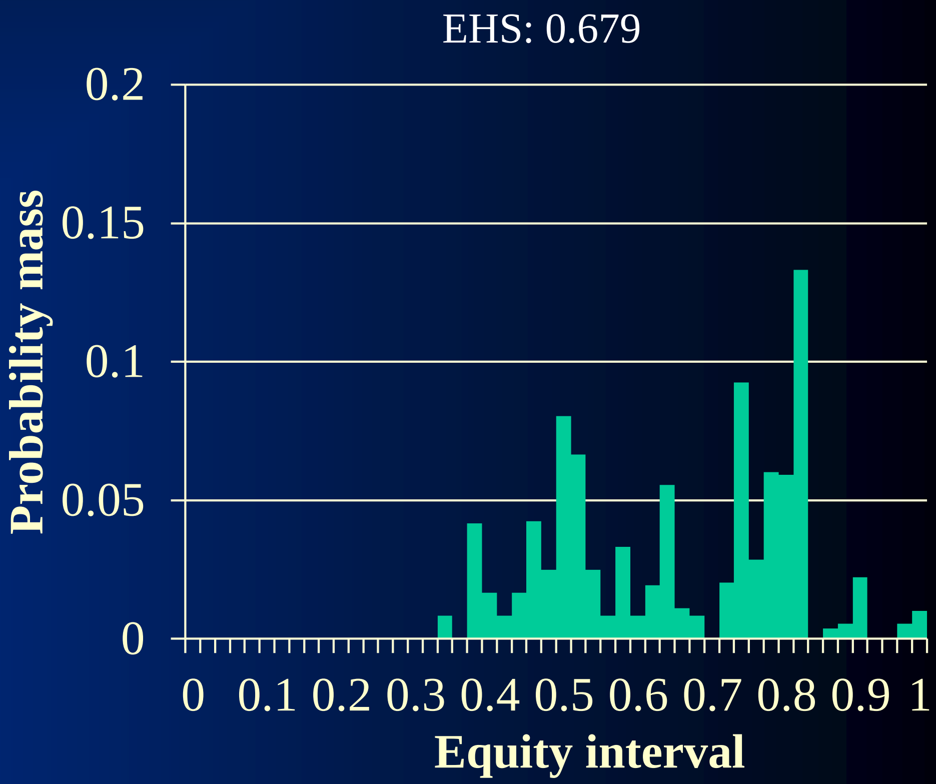


Potential-aware abstraction with EMD

Equity distribution for TcQd-7h9hQh on river (final round)

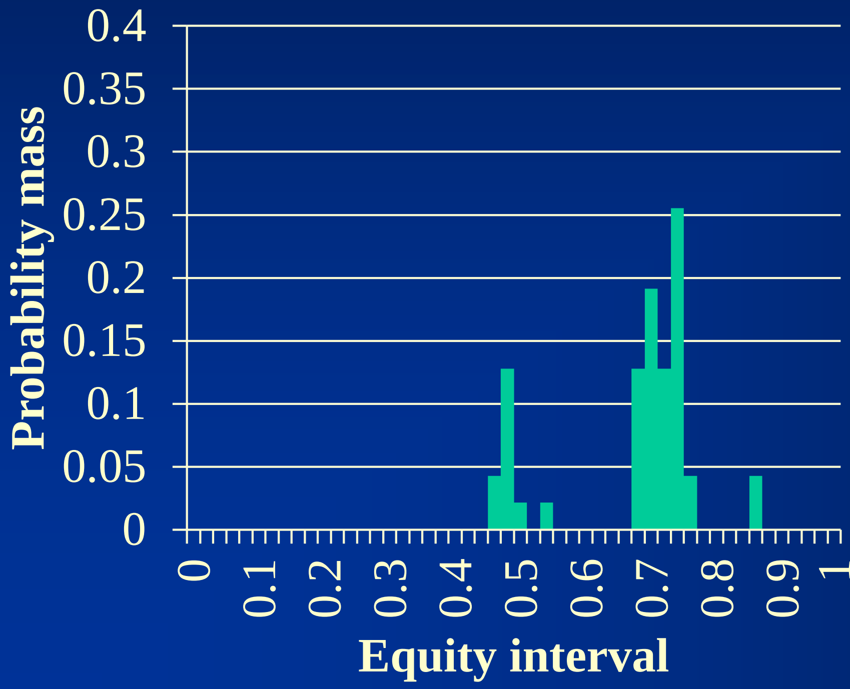


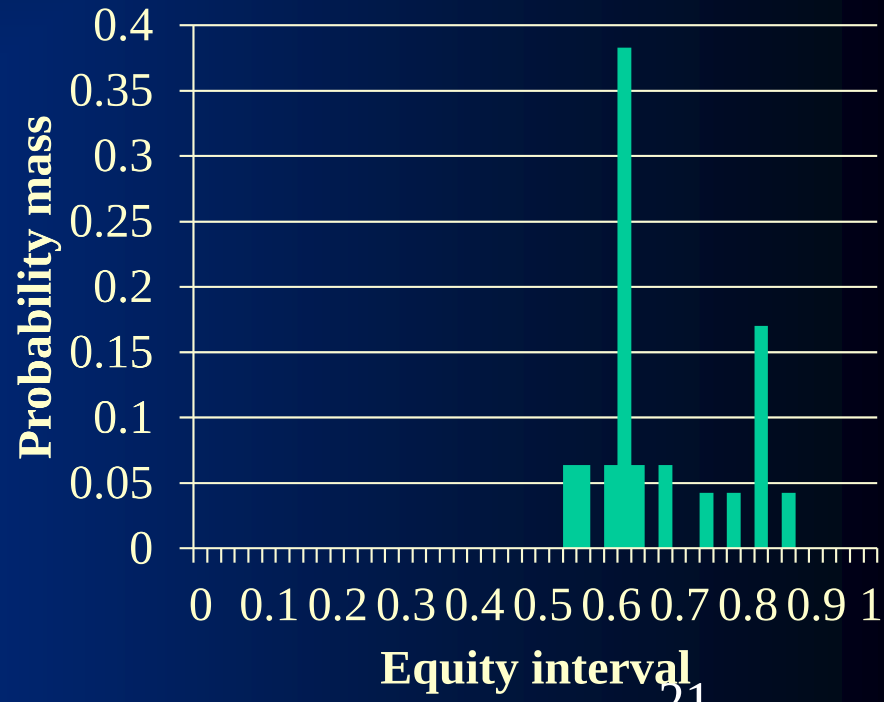
Equity distribution for 5c9d-3d5d7d on river (final round)



Potential-aware abstraction with EMD

- Equity distributions on the turn. Each point is EHS for given turn card assuming uniform random river and opponent hand
- EMD is 4.519 (vs. 0.559 using comparable units to river EMD)

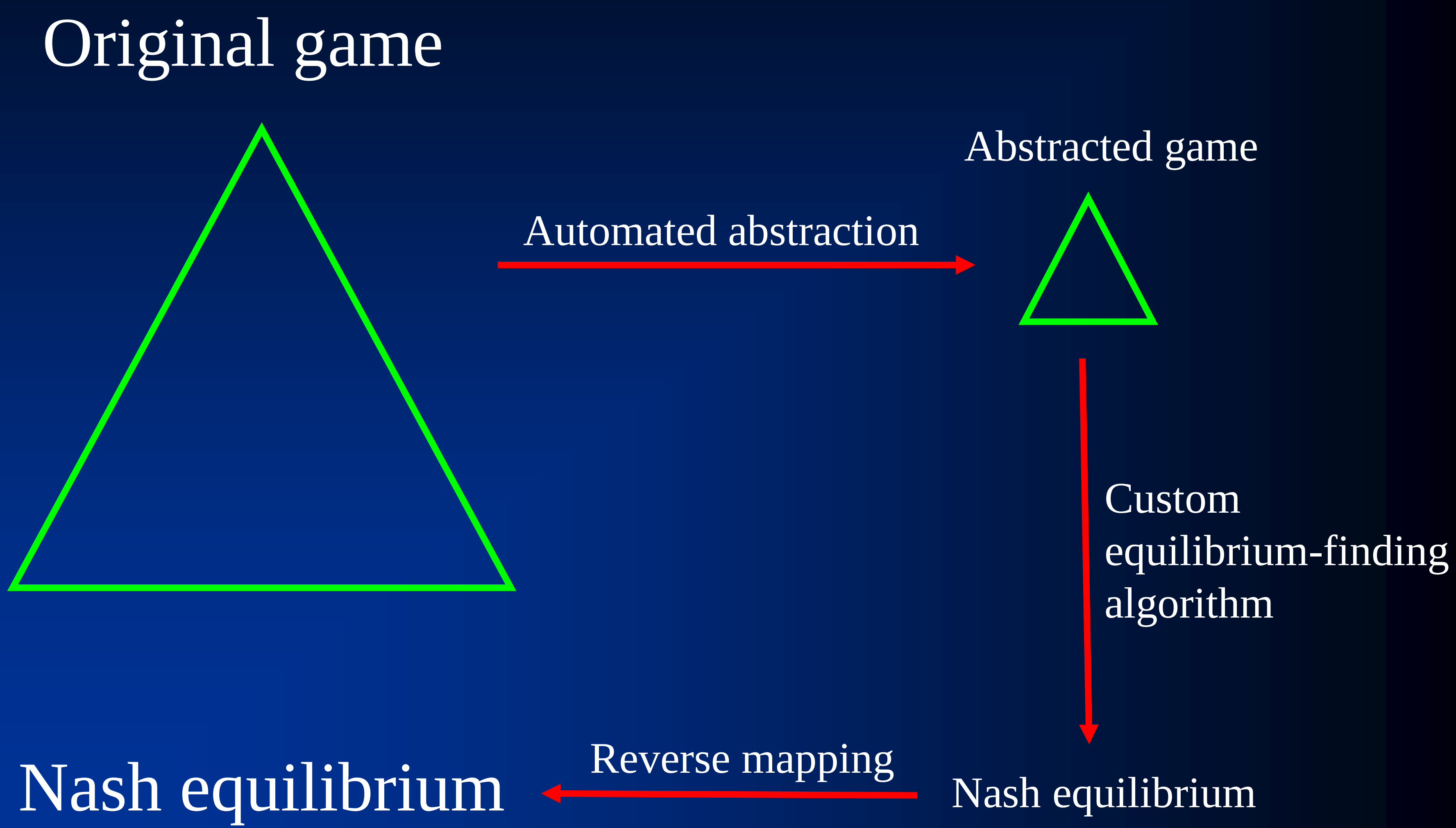




Algorithm for potential-aware imperfect-recall abstraction with EMD

- Bottom-up pass of the information tree (assume an abstraction for final rounds has already been computed using arbitrary approach)
- For each round n
  - Let m^{n+1}_i denote mean of cluster i in A^{n+1}
  - For each pair of round n+1 clusters (i,j), compute distance $d^n_{i,j}$ between m^{n+1}_i and m^{n+1}_j using d^{n+1}
  - For each point x^n, create histogram over clusters from A^{n+1}
  - Compute abstraction A^n using EMD with $d^n_{i,j}$ as ground distance function
    - Developed fast custom heuristic for approximating EMD in our multidimensional setting
    - Best commercially-available algorithm was far too slow to compute abstractions in poker

Standard paradigm for solving large extensive-form games



Hierarchical abstraction to enable distributed equilibrium computation

- On distributed architectures and supercomputers with high inter-blade memory access latency, straightforward MCCFR parallelization approaches lead to impractically slow runtimes
  - When a core does an update at an information set it needs to read and write memory with high latency
  - Different cores working on same information set may need to lock memory, wait for each other, possibly over-write each others' parallel work, and work on out-of-sync inputs
- Our approach solves the former problem and also helps mitigate the latter issue

High-level approach

- To obtain these benefits, our algorithm creates an information abstraction that allows us to assign disjoint components of the game tree to different blades so the trajectory of each sample only accesses information sets located on the same blade.
  - First cluster public information at some early point in the game (public flop cards in poker), then cluster private information separately for each public cluster.
- Run modified version of external-sampling MCCFR
  - Samples one pair of preflop hands per iteration. For the later betting rounds, each blade samples public cards from its public cluster and performs MCCFR within each cluster.

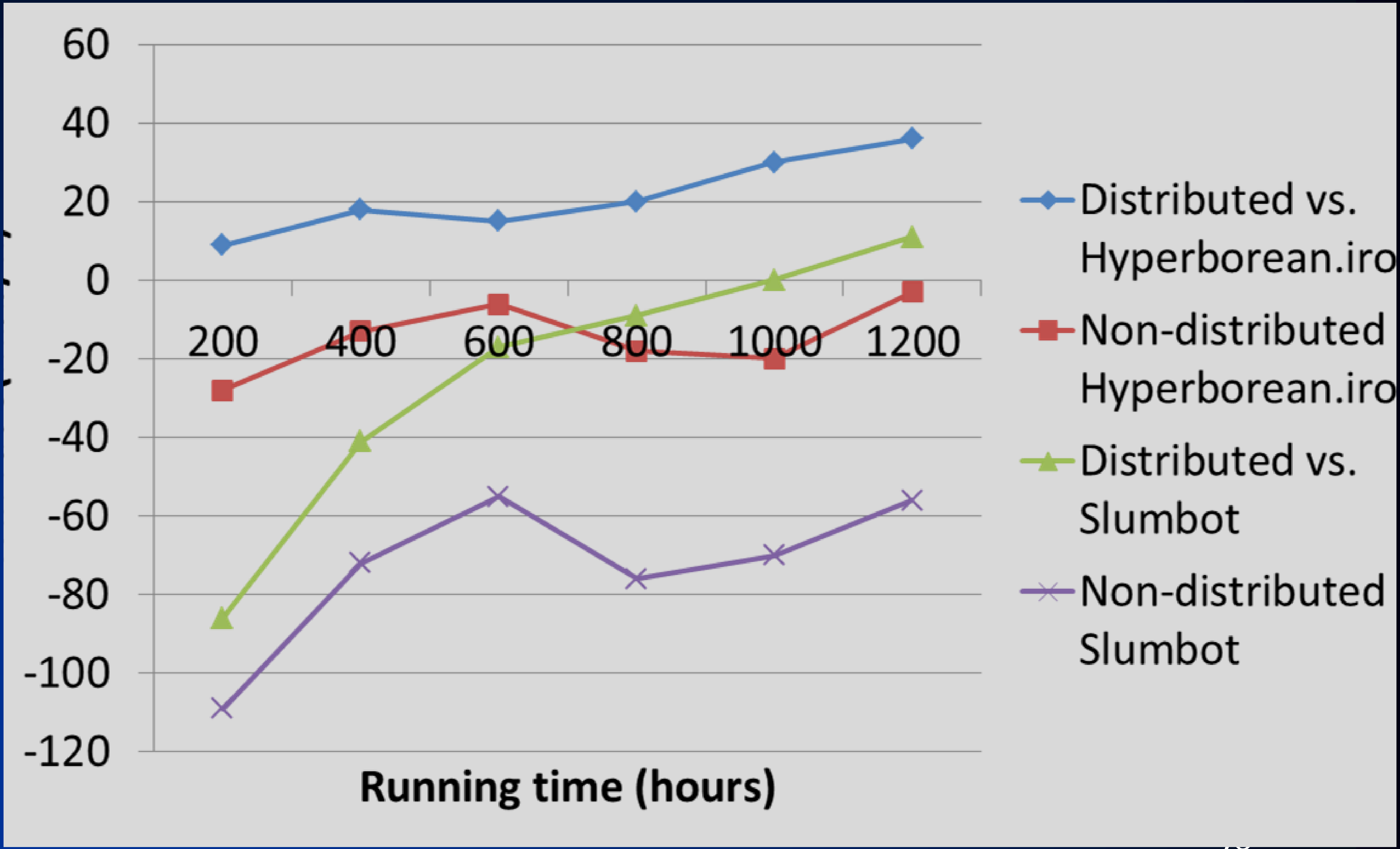
Hierarchical abstraction algorithm for distributed equilibrium computation

- For r = 1 to r*-1, cluster states at round r using A_r
  - A_r is arbitrary abstraction algorithm
  - E.g., for preflop round in poker
- Cluster public states at round r* into C buckets
  - E.g., flop round in poker
- For r = r* to R, c = 1 to C, cluster states at round r that have public information states in public bucket c into B_r buckets using abstraction algorithm A_r

Algorithm for computing public information abstraction

- Construct transition table T
  - T[p][b] stores how often public state p will lead to bucket b of the *base abstraction* A, aggregated over all possible states of private information.
- for i = 1 to M-1, j = i+1 to M (M is # of public states)
  - $s_{i,j} := 0$
  - for b = 1 to B
    - $s_{i,j}$ += min(T[i][b],T[j][b])
  - $d_{i,j} = (V - s_{i,j})/V$
- Cluster public states into C clusters using (custom) clustering algorithm L with distance function d
  - $d_{i,j}$ corresponds to fraction of private states not mapped to same bucket of A when paired with public info i and j

Comparison to non-distributed approach



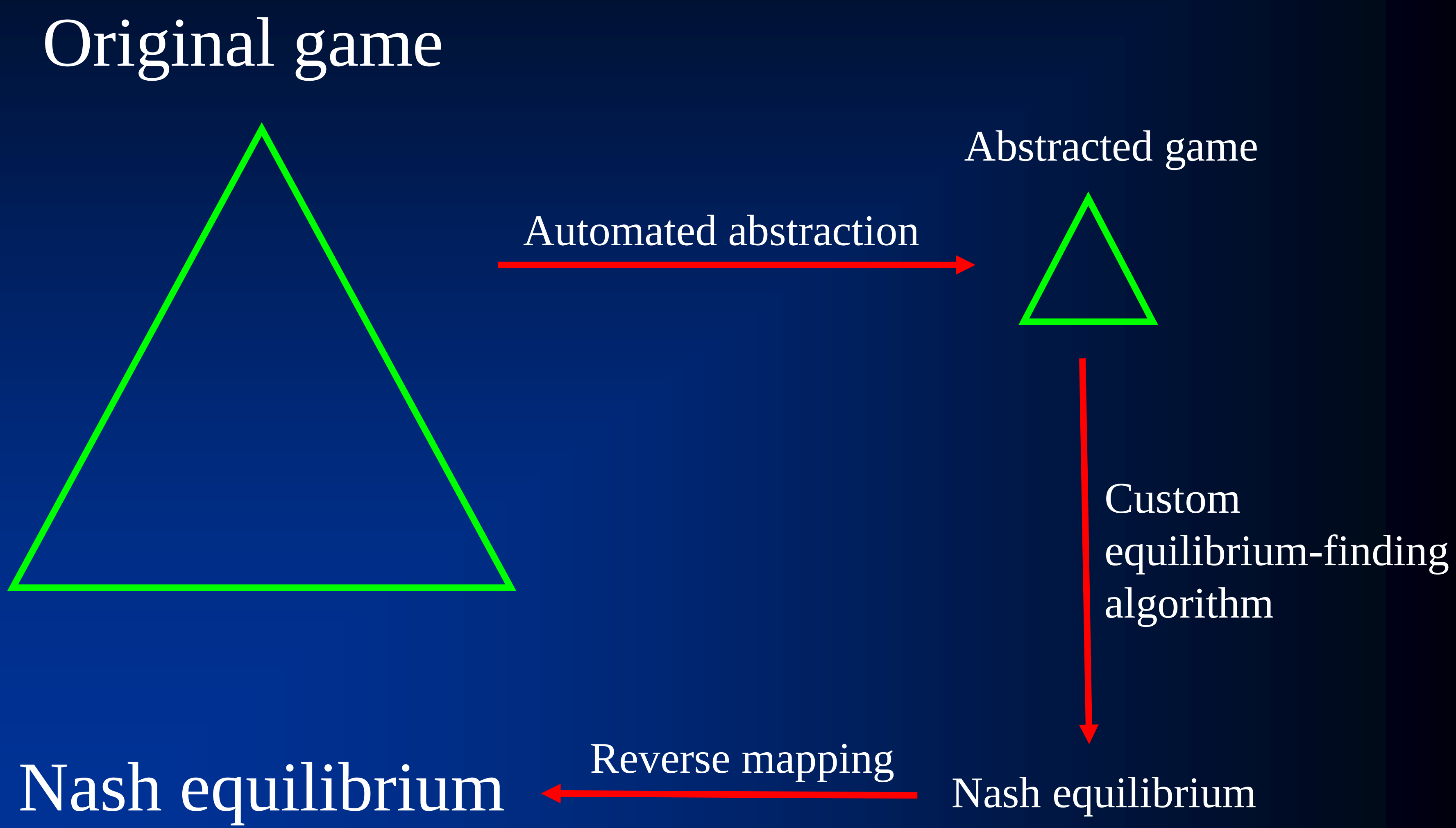
Tartanian7: champion two-player no-limit Texas Hold 'em agent

- Beat every opponent with statistical significance in 2014 AAAI computer poker competition

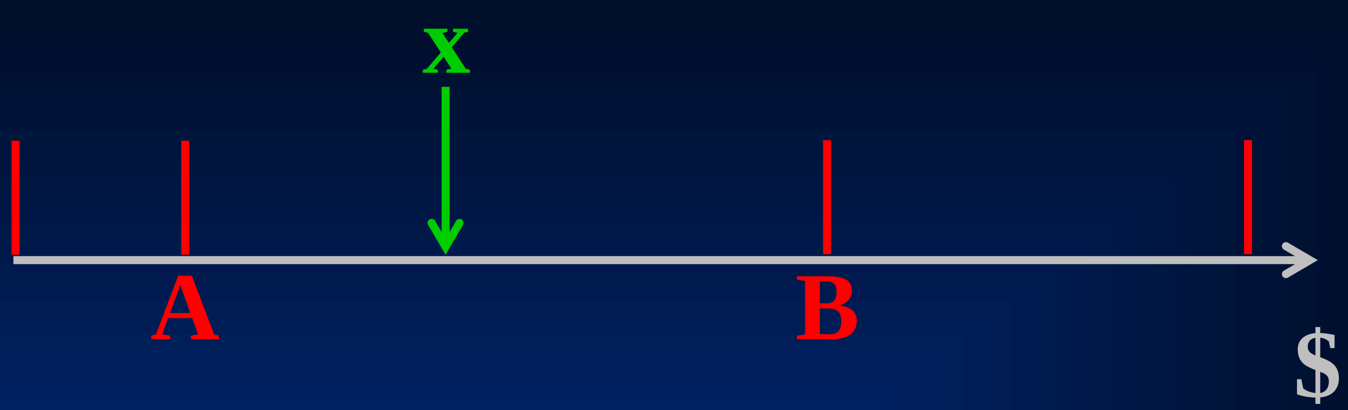
SartreNLExp	Nyx	Hyperborean.iro	Slumbot	Prelude	HibiscusBiscuit	PijaiBot	Feste.iro	LittleRock	KEmpfer	Rembrant3	HITSZ_CS_14	Lucifer
261 ± 47	121 ± 38	21 ± 16	33 ± 16	20 ± 16	125 ± 44	499 ± 68	141 ± 45	214 ± 57	516 ± 61	980 ± 34	1474 ± 180	1819 ± 111

Table 1: Win rate (in mbb/h) of our agent in the 2014 AAAI Annual Computer Poker Competition against opposing agents.

Standard paradigm for solving large imperfect-information games



Action translation



- $f_{A,B}(x) \equiv$ probability we map x to A
  - Will also denote as just $f(x)$

[Ganzfried & Sandholm IJCAI-13]

A natural approach

- If x < $\frac{A+B}{2}$, then map x to A; otherwise, map x to B
- Called the **deterministic arithmetic** mapping

- Suppose pot is 1, stacks are 100
- Suppose we are using the {fold, call, pot, all-in} action abstraction
  - "previous expert knowledge [has] dictated that if only a single bet size [in addition to all-in] is used everywhere, it should be pot sized" [Hawkin et al., AAAI 2012]
- Suppose opponent bets x in (1,100)
  - So A = 1, B = 100

- Suppose we call a bet of 1 with probability ½ with a medium-strength hand
- Suppose the opponent has a very strong hand
- His expected payoff of betting 1 will be:

  $(1 \cdot \frac{1}{2}) + (2 \cdot \frac{1}{2}) = 1.5$
- If instead he bets 50, his expected payoff will be:

  $(1 \cdot \frac{1}{2}) + (51 \cdot \frac{1}{2}) = 26$
- He gains $24.50 by exploiting our translation mapping!
- Tartanian1 lost to an agent that didn’t look at its private cards in 2007 ACPC using this mapping!

An improvement

- What if we randomize and map x to A with probability $\frac{B-x}{B-A}$?
- Suppose opponent bets 50.5, and we call an all-in bet with probability $\frac{1}{101}$ with a mediocre hand
- Then his expected payoff is $13.875
- An improvement, but still way too high
- Called the **randomized arithmetic** mapping

Other prior approaches

- Deterministic geometric: If $\frac{A}{x} > \frac{x}{B}$, map x to A; otherwise, map x to B
  - Used by Tartanian2 in 2008
- Randomized geometric 1
  - $f(x) = \frac{A(B-x)}{A(B-x) + x(x-A)}$
  - Used by Alberta 2009-present
- Randomized geometric 2
  - $f(x) = \frac{A(B+x)(B-x)}{(B-A)(x^2 + AB)}$
  - Used by CMU 2010-2011

Problem with prior approaches?

- High exploitability in simplified games
- Purely heuristic and not based on any theoretical justification
- Fail to satisfy natural theoretical properties

Our new mapping

- We propose a new mapping, called the pseudo-harmonic mapping, which is the only mapping consistent with the equilibrium of a simplified poker game:
  - $f(x) = \frac{(B-x)(1+A)}{(B-A)(1+x)}$
- This mapping has significantly lower exploitability than the prior ones in several simplified poker games
- Significantly outperforms the randomized-geometric mappings in no-limit Texas hold'em

Action translation desiderata

1. Boundary constraints: $f(A) = 1$, $f(B) = 0$
2. Monotonicity
3. Scale invariance
4. Action robustness: small change in x doesn't lead to large change in f
5. Boundary robustness: small change in A or B doesn't lead to large change in f

Theoretical results

- Randomized geometric mappings violate boundary robustness. If we allow $A = 0$ they are discontinuous in A. Otherwise, they are Lipschitz-discontinuous in A.
- Only randomized-arithmetic and randomized-pseudo-harmonic satisfy all the desiderata

Purification and thresholding

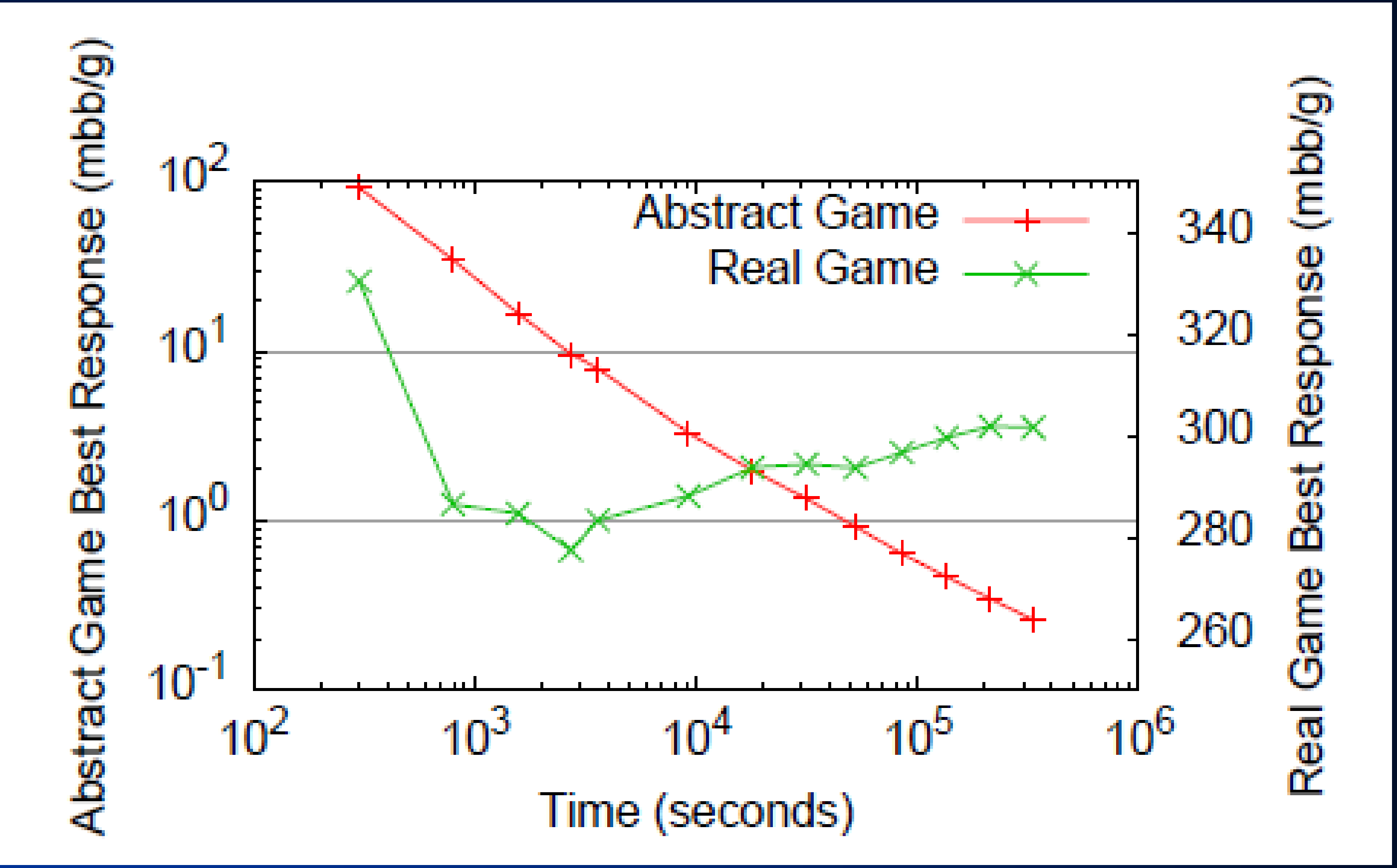
- *Thresholding*: round action probabilities below c down to 0 (then renormalize)
- *Purification* is extreme case where we play maximal-probability action with probability 1

Benefits of post-processing techniques

- 1) Failure of equilibrium-finding algorithm to fully converge
  – Tartanian4 had exploitability of 800 mbb/hand even within its abstraction (always folding has exploitability of 750 mbb/hand!)

Benefits of post-processing techniques

- 2) Combat overfitting of equilibrium to the abstraction



Acknowledgement: Johanson et al., AAAI ‘12

Experiments on no-limit Texas hold ‘em

- Purification outperforms using a threshold of 0.15
  - Does better than it against all but one 2010 competitor, beats it head-to-head, and won bankroll competition

Worst-case exploitability

- We also compared *worst-case exploitabilities* of several variants submitted to the 2010 two-player limit Texas hold 'em division
  - Using algorithm of Johanson *et al.* IJCAI-11

Threshold	Exploitability of GS6	Exploitability of Hyperborean
None	463.591	**235.209**
0.05	326.119	243.705
0.15	**318.465**	258.53
0.25	335.048	277.841
Purified	349.873	437.242

Table 4: Results for full-game worst-case exploitabilities of several strategies in two-player limit Texas Hold'em. Results are in milli big blinds per hand. Bolded values indicate the lowest exploitability achieved for each strategy.

Purification and thresholding

- 4x4 two-player zero-sum matrix games with payoffs uniformly at random from [-1,1]
- Compute equilibrium F in full game
- Compute equilibrium A in abstracted game that omits last row and column
  - essentially "random" abstractions
- Compare $u_1(A_1, F_2)$ to $u_1(\mathrm{pur}(A_1), F_2)$
- **Conclusion: Abstraction+purification outperforms just abstraction (against full equilibrium) at 95% confidence level**

Purification and thresholding

Purified average payoff	-0.050987 +- 0.00042
Unpurified average payoff	-0.054905 +- 0.00044
# games where purification led to improved performance	261569 (17.44 %)
# games where purification led to worse performance	172164 (11.48%)
# games where purification led to no change in performance	1066267 (71.08 %)

- Some conditions when they perform identically:
  1. The abstract equilibrium A is a pure strategy profile
  2. The support of A_1 is a subset of the support of F_1

Purification and thresholding

- Results depend crucially on the support of the full equilibrium
- If we only consider the set of games that have an equilibrium σ with a given support, purification improves performance for each class except for the following, where the performance is statistically indistinguishable:
  - σ is the pure strategy profile in which each player plays his fourth pure strategy
  - σ is a mixed strategy profile in which player 1's support contains his fourth pure strategy, and player 2's support does not contain his fourth pure strategy

New family of post-processing techniques

- 2 main ideas:
  - Bundle similar actions
  - Add preference for conservative actions
- First separate actions into {fold, call, “bet”}
  - If probability of folding exceeds a threshold parameter, fold with prob. 1
  - Else, follow purification between fold, call, and “meta-action” of “bet.”
  - If “bet” is selected, then follow purification within the specific bet actions.
- Many variations: threshold parameter, bucketing of actions, thresholding value among buckets, etc.

Post-processing experiments

	Hyperborean.iro	Slumbot	Average	Min
No Thresholding	+30 ± 32	+10 ± 27	+20	+10
Purification	+55 ± 27	+19 ± 22	+37	+19
Thresholding-0.15	+35 ± 30	+19 ± 25	+27	+19
New-0.2	+39 ± 26	+103 ± 21	+71	+39

Brains vs. Artificial Intelligence

- April 24-May 8, 2015 at Rivers Casino in Pittsburgh, PA
  - The competition was organized by Carnegie Mellon University Professor Tuomas Sandholm. Collaborators were Tuomas Sandholm and Noam Brown.
- 20,000 hands of two-player no-limit Texas hold ‘em between “Claudico” and Dong Kim, Jason Les, Bjorn Li, Doug Polk
  - 80,000 hands in total
  - Two 750-hand sessions per day

“Duplicate” scoring

- Suppose Dong has pocket aces and Claudico has pocket kings and Dong wins \$5,000. Did Dong “outplay” Claudico?
  - What if Bjorn had pocket kings against Claudico’s pocket aces in the same situation (same public cards dealt on the board), and Claudico won \$10,000?

Brains

Doug Polk
@DougPolkPoker

March HUNL PR
1 West Coast Gangsters
2 Big Dick
3 AZNflushie (RIP)
4 Rumble man
5 Swarmmy
6 Kaby
7 Ike
8 wheyprotein
9 80%carry
10 muumi

RETWEETS 4 LIKES 13

4:24 AM - 28 Feb 2015

Nick Frame
@TCfromUB
Follow

The REAL power rankings for OCT 2014 are out

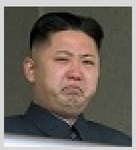
TC power rankings OCT 2014

1. WCG (0)
2. ike (+1)
3. sauce (+1)
4. TCfromUB (+1)
5. jungle (+5)
6. pandorasbux (-4)
7. kabydf (0)
8. donger (-2)
9. carrycakes (-1)
10. KPR (-1)
11. asianflushie (+3)
12. kanu7 (+3)
13. bajskorven (U)
14. OTBredbaron (U)
15. Rperfumo (-4)
16. mokoma1 (0)
17. Billiomucks (-5)
18. dougiedan (-5)
19. ForTheSwarm (U)
20. Willhasha (U)

Brains

02-04-2015, 09:53 AM #1

Donger Kim
enthusiast



Join Date: Feb 2015
Location: South Korea
@dongerkim
Posts: 54

Donger Kim to Nick Frame (TCfromUB) HU Challenge

I am a high-stakes heads up nlhe regular on PokerStars where I play under the name "Donger Kim". There's been quite a bit of discussion on heads-up rankings lately, particularly from TCfromUB (Nick Frame, TooCuriosso1 on 2p2). I've played quite a bit with him and think he's a top player. I respect his game and it would be humbling to play him and represent my country.

However, as he ranks himself ahead of me, I'd like to have a chance to play him in a challenge-type format. I think it would be a fun experience and something that would also be enjoyable for the community.

I propose we do a 15k hand challenge at 100/200 nl with a $50k sidebet escrowed with ike or sauce. I suggest we put some reasonable time frame conditions on this, we're both grinders so we should be able to finish this in a 1-2 week time frame.

Nick, let me know when you'd like to begin. Ideally, I'd like to get started right away.

Brains

MAR 2

Donger Kim wins heads-up challenge against TCfromUB

Poker News

G+1 0 Like 1 Share

Dong "Donger Kim" Kim won $103.992 from Nick "TCfromUB" Frame in the15,000 hand heads-up challenge, which not only earned him the respect of the high stakes community, but also an additional $15,000 from the sidebets for the challenge.

Results

- Humans won by 732,713 chips, which corresponds to 9.16 big blinds per 100 hands (BB/100) (SB = 50, BB = 100). Players started each hand with 200 big blinds (20,000 chips).
  - Statistically significant at 90% confidence level, but not 95%
- Dong Kim beat Nick Frame by 13.87 BB/100
  - $103,992 over 15,000 hands with 25-50 blinds
- Doug Polk beat Ben Sulsky by 24.67 BB/100
  - $740,000 over 15,000 hands with 100-200 blinds

Payoffs

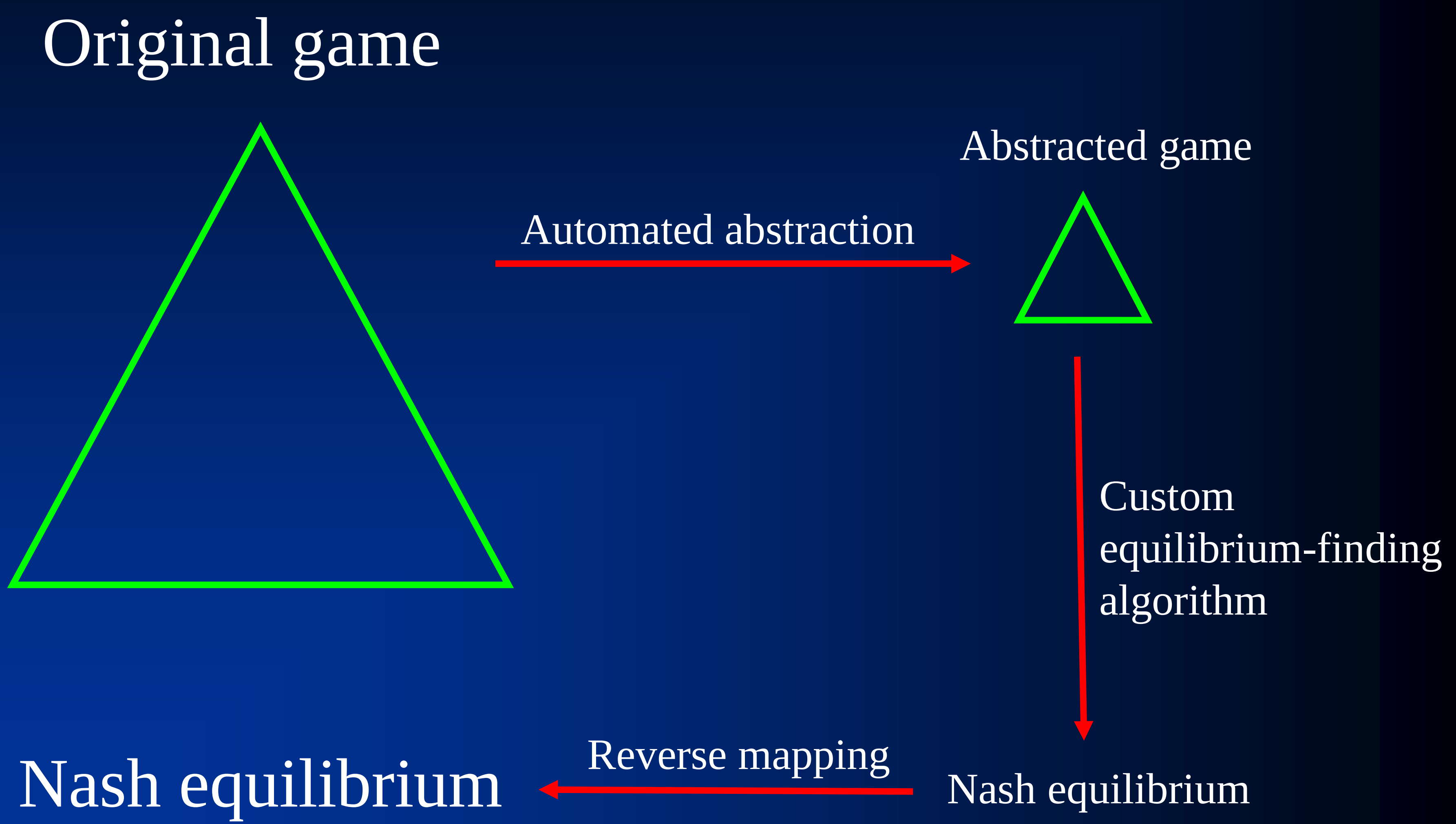
- Prize pool of \$100,000 distributed to the humans depending on their individual profits.

$$
\begin{aligned}
\text{If } x_1 &> x_4 \\
p_1 &= \$10,000 + \$60,000 \cdot \frac{x_1 - x_4}{x_1 + x_2 + x_3 - 3x_4} \\
p_2 &= \$10,000 + \$60,000 \cdot \frac{x_2 - x_4}{x_1 + x_2 + x_3 - 3x_4} \\
p_3 &= \$10,000 + \$60,000 \cdot \frac{x_3 - x_4}{x_1 + x_2 + x_3 - 3x_4} \\
p_4 &= \$10,000 \\
\text{Else} & \\
p_1 &= p_2 = p_3 = p_4 = \$25,000
\end{aligned}
$$

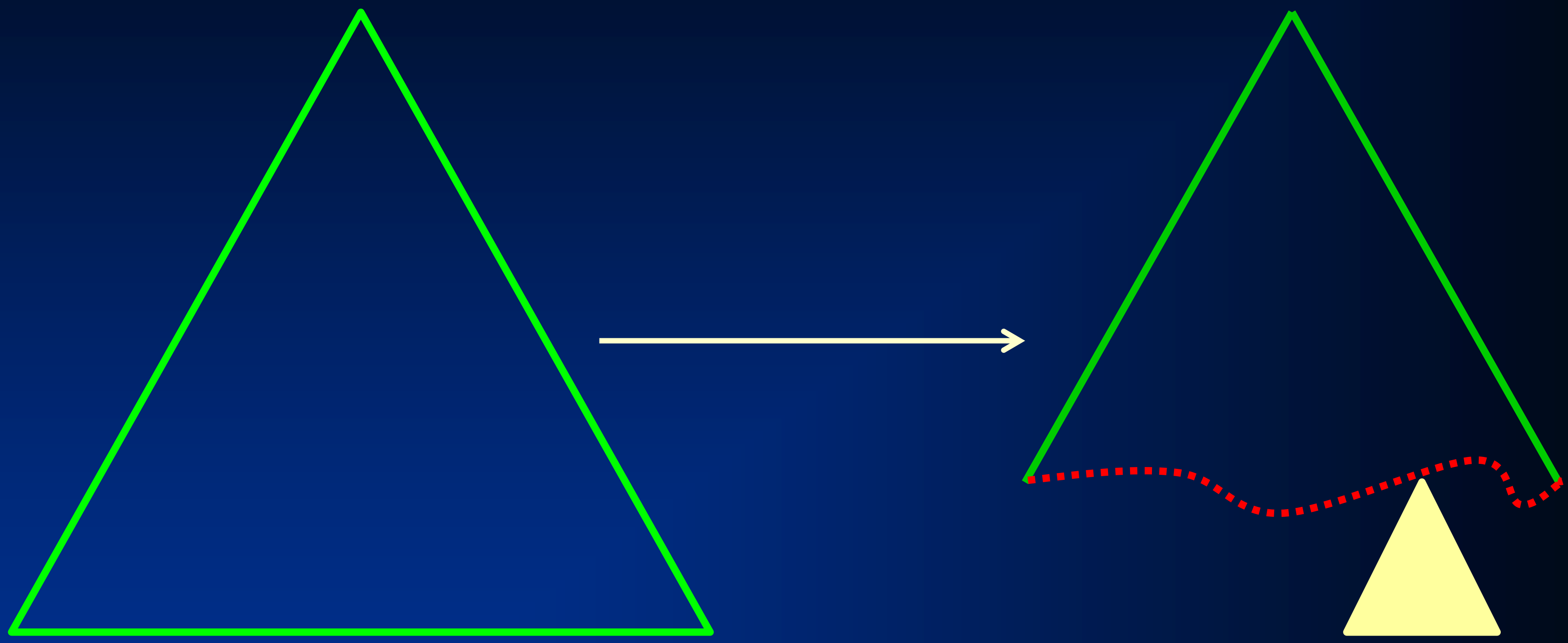
I Limp!

- "**Limping is for Losers**. This is *the most important fundamental* in poker -- for every game, for every tournament, every stake: If you are the first player to voluntarily commit chips to the pot, open for a raise. Limping is inevitably a losing play. If you see a person at the table limping, you can be fairly sure he is a bad player. Bottom line: If your hand is worth playing, it is worth raising" [Phil Gordon's Little Gold Book, 2011]
- Claudico limps close to 10% of its hands
  - Based on humans' analysis it profited overall from the limps
- Claudico makes many other unconventional plays (e.g., small bets of 10% pot and all-in bets for 40 times pot)

Standard paradigm for solving large imperfect-information games



Endgame solving



Strategies for entire game computed offline

Endgame strategies computed in real time to greater degree of accuracy

Endgame definition

- E is an **endgame** of a game G if:

1. Set of E's nodes is a subset of set of G's nodes
2. If s' is a child of s in G and s is a node in E, then s' is also a node in E
3. If s is in the same information set as s' in G and s is a node in E, then s' is also a node in E

Can endgame solving guarantee equilibrium?

- Suppose that we computed an exact (full-game) equilibrium in the initial portion of the game tree prior to the endgame (the **trunk**), and computed an exact equilibrium in the endgame. Is the combined strategy an equilibrium of the full game?

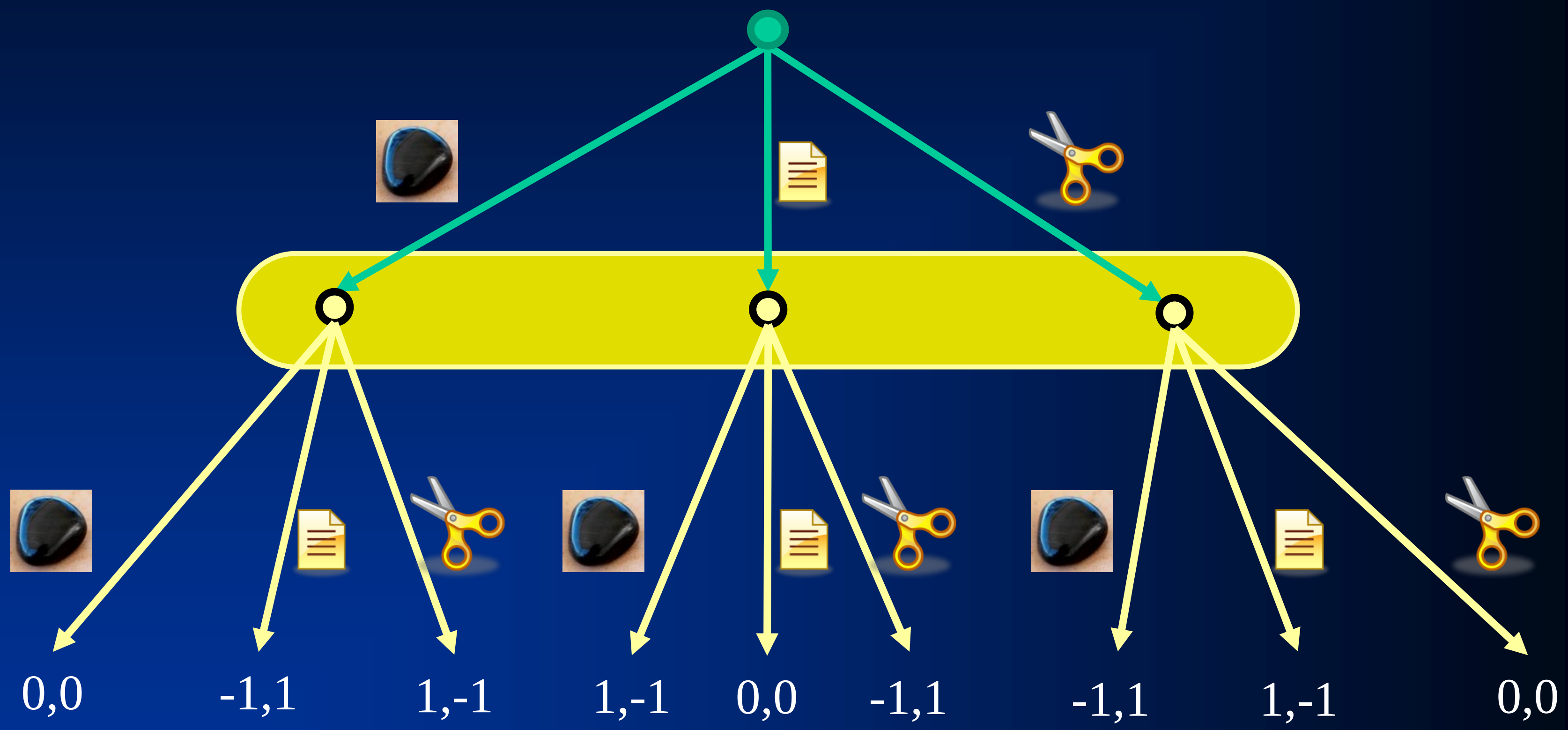
Can endgame solving guarantee equilibrium?

- No!
- Several possible reasons this may fail:
  - The game may have many equilibria, and we might choose one for the trunk that does not match up correctly with the one for the endgame
  - We may compute equilibria in different endgames that do not balance appropriately with each other

Can endgame solving guarantee equilibrium?

Proposition: There exist games with a unique equilibrium and a single endgame for which endgame solving can produce a non-equilibrium strategy profile in the full game

Limitation of endgame solving



Benefits of endgame solving

- Computation of exact (rather than approximate) equilibrium strategies in the endgames
- Computation of equilibrium refinements (e.g., undominated and ε-quasi-perfect equilibrium)
- Better abstractions in the endgame that is reached
  - Finer-grained abstractions
  - History-aware abstractions
  - Strategy-biased abstractions
- Solving the "off-tree problem"

Efficient algorithm for endgame solving in large imperfect-information games

- Naïve approach requires $O(n^2)$ lookups to the strategy table, where n is the number of possible hands
  - Computationally infeasible (> 1 min/hand)
- Our algorithm uses just O(n) strategy table lookups (8 seconds/hand using Gurobi's LP solver)
- Our approach improved performance against strongest 2013 ACPC agents
  - 87+-50 vs. Hyperborean and 29+-25 vs. Slumbot

- New game decomposition approach (CFR-d) has theoretical guarantee but performs worse empirically
  - Burch et al. AAAI-14
- Recent approach for safer endgame solving that maximizes the "endgame margin"
  - Moravic et al. AAAI-16
- Doug Polk related to me in personal communication after the competition that he thought the river strategy of Claudico using the endgame solver was the strongest part of the agent.

Problematic hands

1. We had A4s and folded preflop after putting in over half of our stack (human had 99).
   - We only need to win 25% of time against opponent's distribution for call to be profitable (we win 33% of time against 99).
   - Translation mapped opponent's raise to smaller size, which caused us to look up strategy computed thinking that pot size was much smaller than it was (7,000 vs. 10,000)
2. We had KT and folded to an all-in bet on turn after putting in ¾ of our stack despite having top pair and a flush draw
   - Human raised slightly below smallest size in our abstraction and we interpreted it as a call
   - Both 1 and 2 due to "off-tree problem"
3. Large all-in bet of 19,000 into small pot of 1700 on river without "blocker"
   - E.g., 3s2c better all-in bluff hand than 3c2c on JsTs4sKcQh
   - Endgame information abstraction algorithm doesn't fully account for "card removal"

Reflections on the First Man vs. Machine No-Limit Texas Hold 'em Competition [Sigecom Exchanges '15, to appear in AI Magazine]

- Two most important avenues for improvement
  - Solving the "off-tree problem"
  - Improved approach for information abstraction that better accounts for card removal/"blockers"
- Improved theoretical understanding of endgame solving
  - Works very well in practice despite lack of guarantees
  - Newer decomposition approach with guarantees does worse
- Bridge abstraction gap
  - Approaches with guarantees only scale to small games
  - Larger abstractions work well despite theoretical "pathologies"
- Diverse applications of equilibrium computation
- Theoretical questions for action translation/post-processing

Second Brains vs. AI Competition

- Libratus: +14.7 BB/100 over 120,000 hands ($200k in prizes)
  - Claudico -9.16 BB/100 over 80,000 hands ($100k in prizes)

1. Libratus: 20-25 million core hours on supercomputer
   – Claudico: 2-3 million core hours on supercomputer

2. Improved equilibrium-finding algorithm "Regret-based pruning" which prunes actions with high regret early on in CFR so that the computation can eliminate large portions of the game tree following these "bad" actions.
   – Brown and Sandholm, "**Reduced Space and Faster Convergence in Imperfect-Information Games via Regret-Based Pruning**," 2017 AAAI Workshop on Computer Poker and Imperfect Information

3. Improved endgame solver. Used supercomputer resources in real time. Was able to solve full turn and river endgames within around 20 seconds. Estimated that it would take 10+ minutes on normal machine.
   - Brown and Sandholm, "**Safe and Nested Endgame Solving in Imperfect-Information Games**," 2017 AAAI Workshop on Computer Poker and Imperfect Information

4. Claudico's mistakes → Libratus' strengths
   - e.g., card removal/"blockers" and off-tree problem

Game solving challenges

- Nash equilibrium lacks theoretical justification in certain game classes
  - E.g., games with more than two players
  - Even in two-player zero-sum games, certain refinements are preferable
- Computing Nash equilibrium is PPAD-complete in certain classes
- Even approximating NE in 2p zero-sum games very challenging in practice for many interesting games
  - Huge state spaces
- Robust exploitation is preferable

Frameworks and directions

- Standard paradigm
  - Abstraction, equilibrium-finding, reverse mapping (action translation and post-processing)
- New paradigms
  - Incorporating qualitative models (can be used to generate human-understandable knowledge)
  - Real-time endgame solving
- Domain-independent approaches
- Approaches are applicable to games with more than two players
  - Direct: abstraction, translation, post-processing, endgame solving, qualitative models, exploitation algorithm
  - Equilibrium algorithms also, but lose guarantees
  - Safe exploitation, but guarantees maxmin instead of value

- www.ganzfriedresearch.com
- https://www.youtube.com/watch?v=phRAyF1rq0I