Regret Minimization in Games and the Development of Champion Multiplayer Computer Poker Agents

University of Alberta
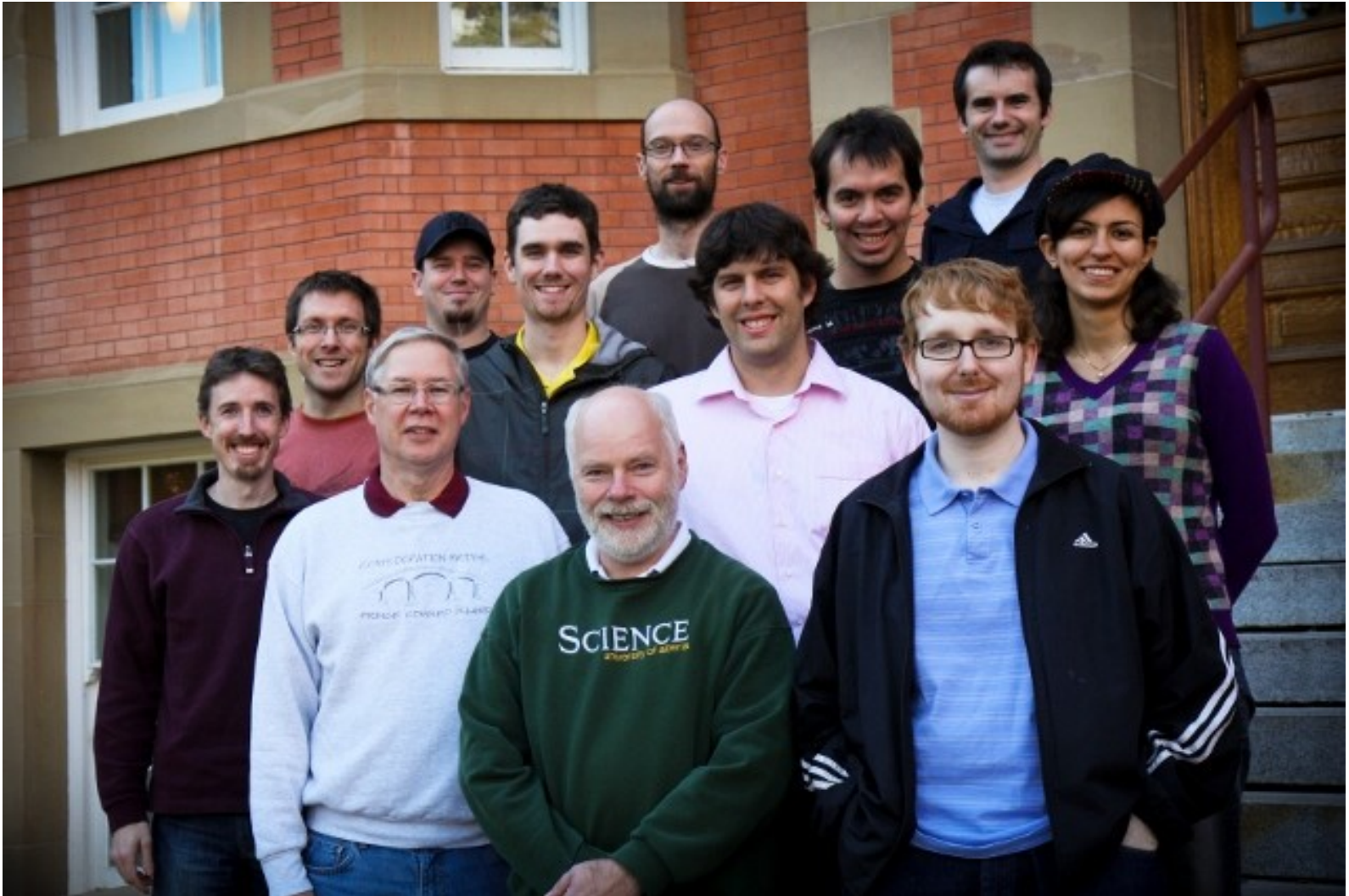Computer Poker Research Group

Richard Gibson

Ph.D. Thesis Presentation
December 6, 2013

Computer Poker Research Group

Heads Up Limit Texas Hold'em
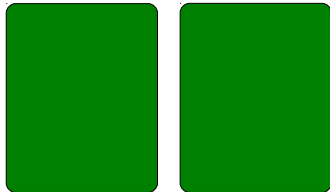
Bet!

Fold?
Call?
Raise?
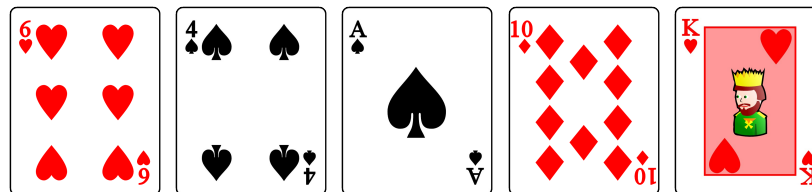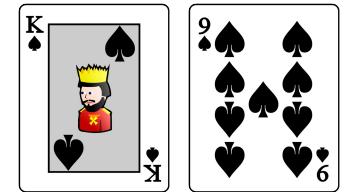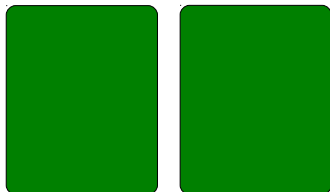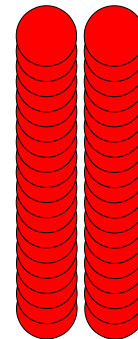
Heads Up No-limit Texas Hold'em

Bet!

All-in!

3-Player Limit Texas Hold'em

Call.

Bet!
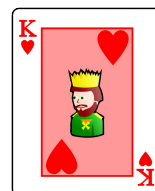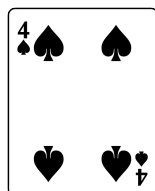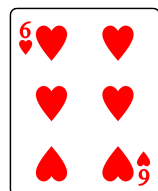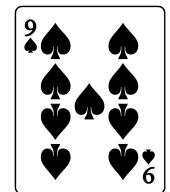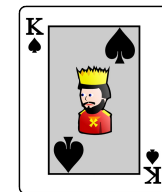
Fold?
Call?
Raise?
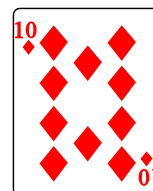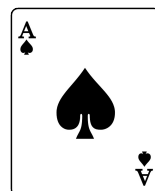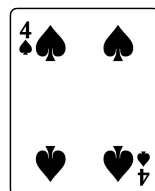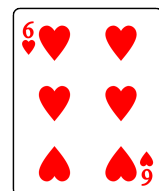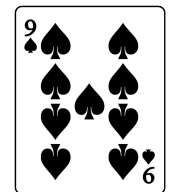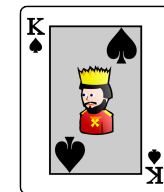
3-Player Limit Texas Hold'em

Source: toonpool.com

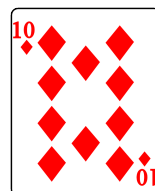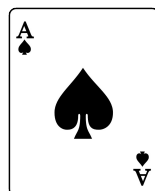2010 - 2013

Hyperborean3p

Hyperborean3p

2009

- No theory
    - 3-player
    - Imperfect recall
- Slow
- Memory expensive

Hyperborean3p

2009

- No theory
  - 3-player
  - Imperfect recall
- Slow
- Memory expensive

2013

- New theory
  - Many players
  - Imperfect recall
- Fast
- Improved performance with limited memory

Outline of Presentation

- Background

  - Counterfactual Regret Minimization (CFR)

- Theoretical Advancements for CFR in:

  - Many player games

  - Imperfect recall games

- CFR Speed-Ups

- Tricks with Memory Limitations

- Conclusion + Future Work

Outline of Presentation

- **Background**
    - Counterfactual Regret Minimization (CFR)
- Theoretical Advancements for CFR in:
    - Many player games
    - Imperfect recall games
- CFR Speed-Ups
- Tricks with Memory Limitations
- Conclusion + Future Work

Background - Kuhn Poker
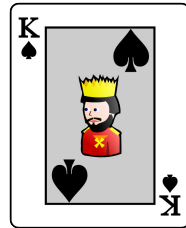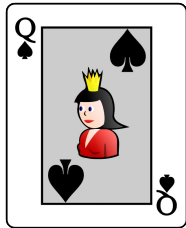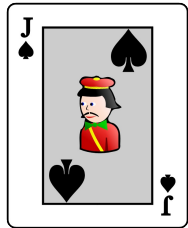
Background - Kuhn Poker
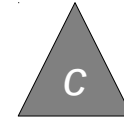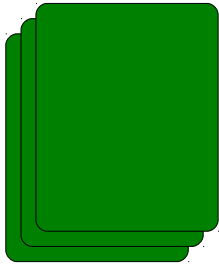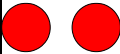
Background - Kuhn Poker
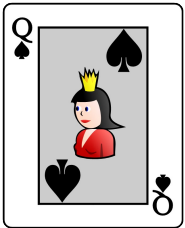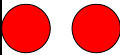
Background - Kuhn Poker

Check / Bet ?
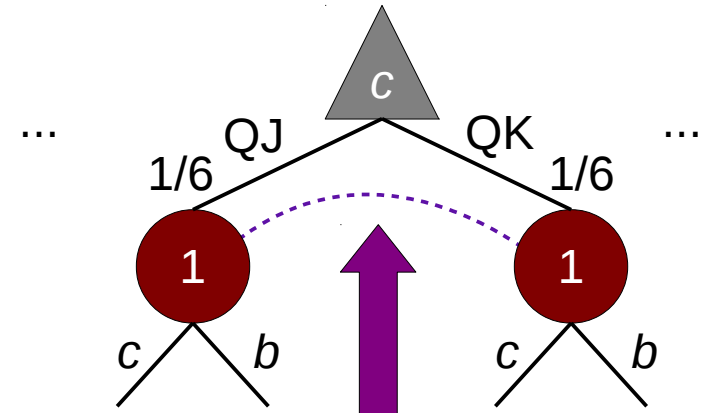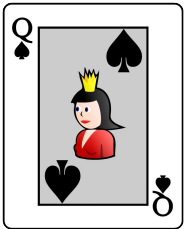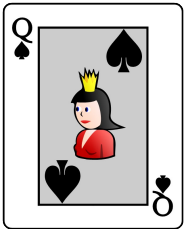
Information set

Background - Kuhn Poker

Bet!

Fold / Call ?

Background - Kuhn Poker
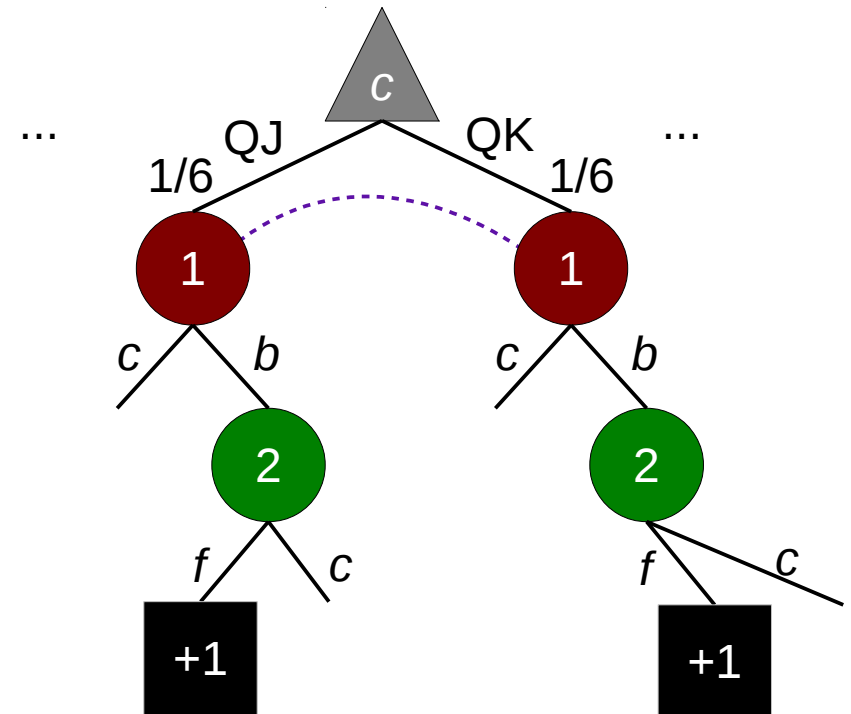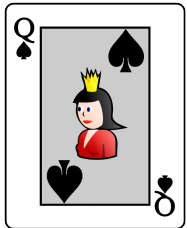
Background - Kuhn Poker
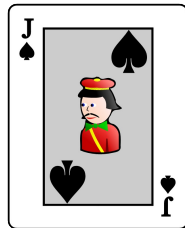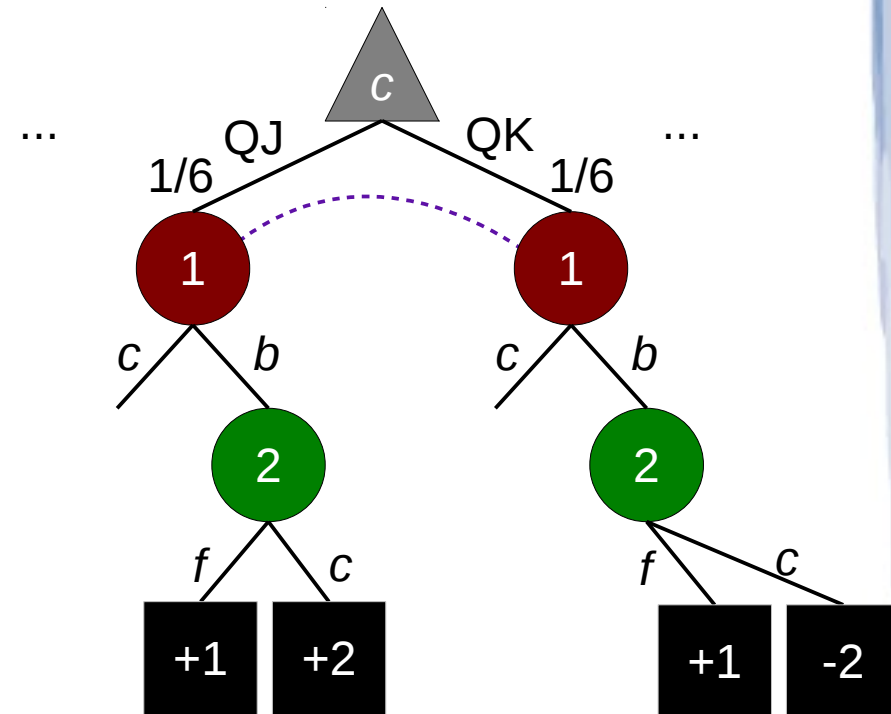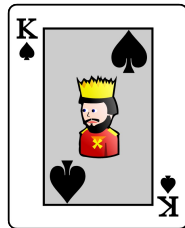
Background - Kuhn Poker
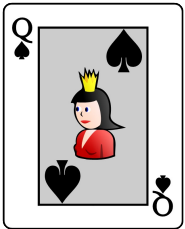
Check.

Check / Bet ?

Background - Kuhn Poker

Check.   Check.
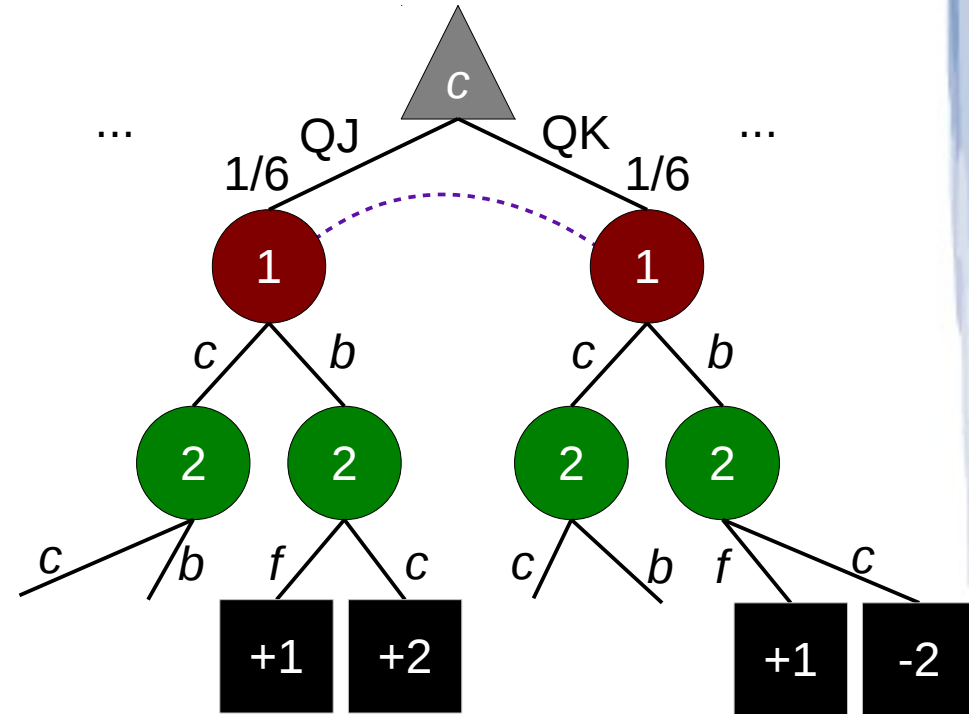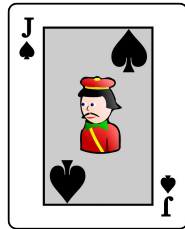
+1 / -1   -1 / +1
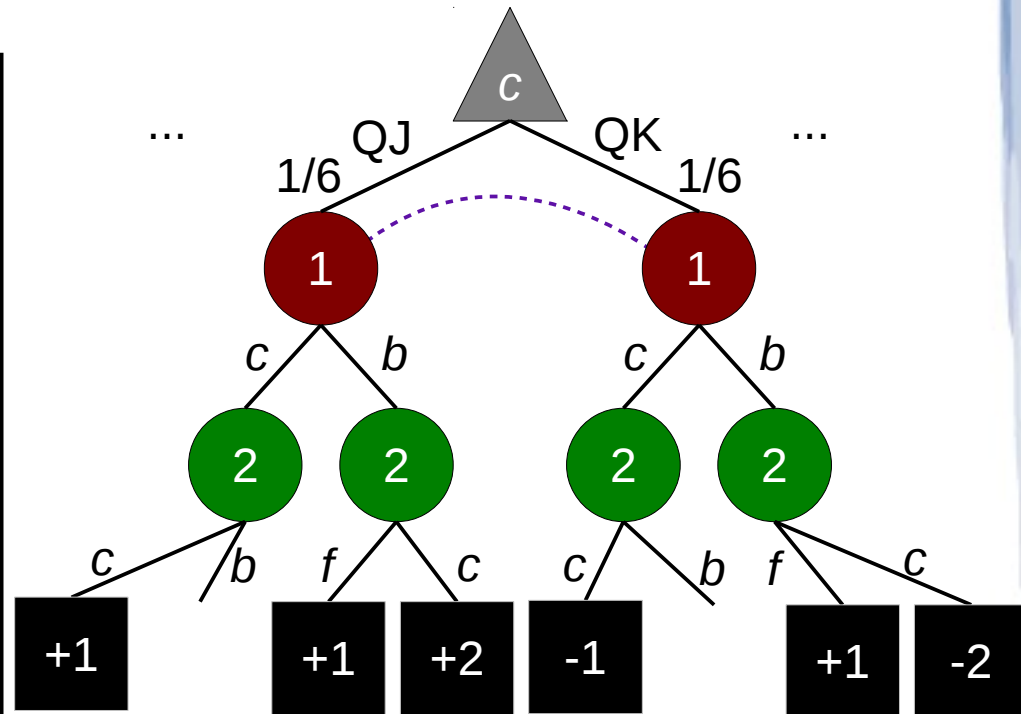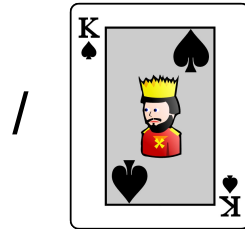
Background - Kuhn Poker

Fold / Call ?   Bet!
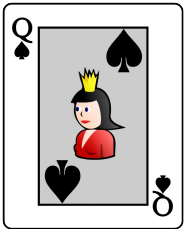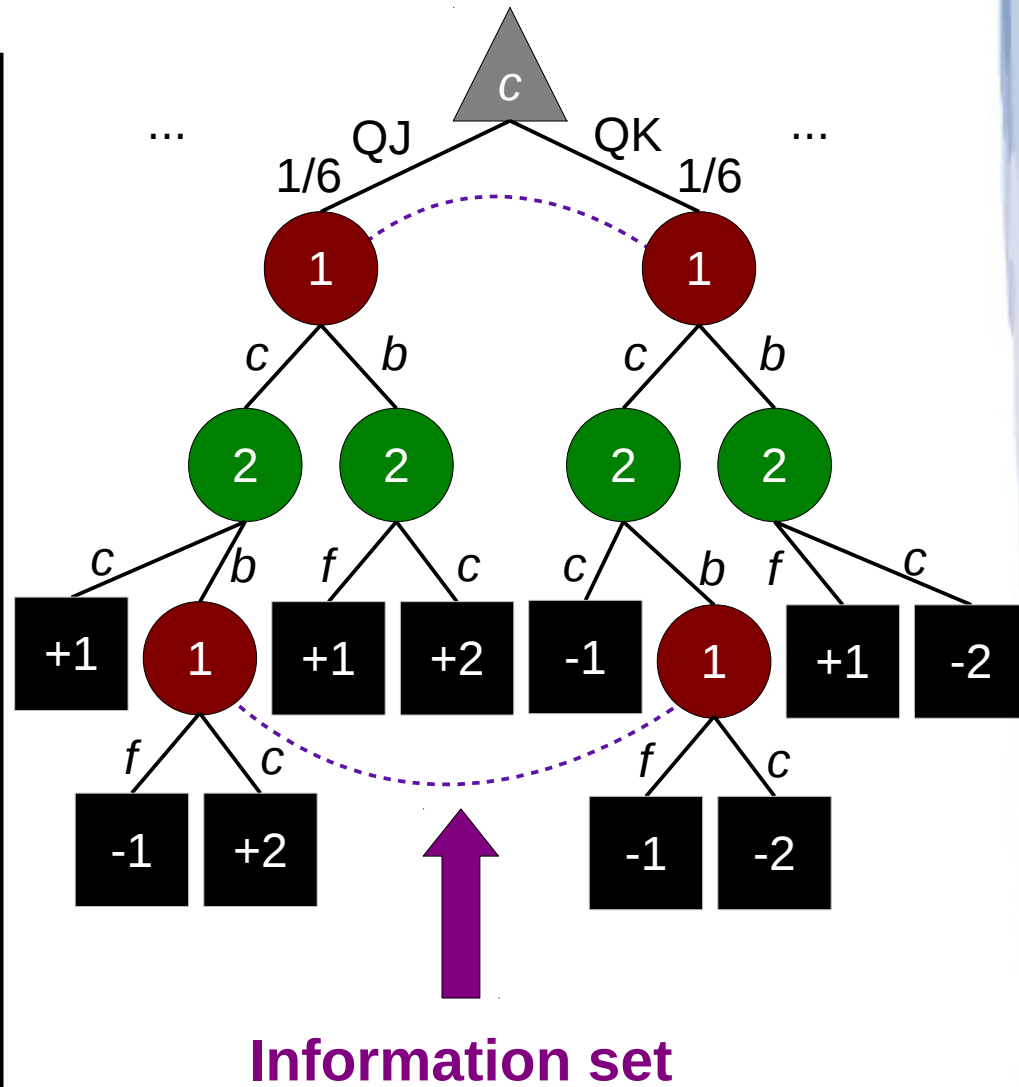
Information set

Background

In general:
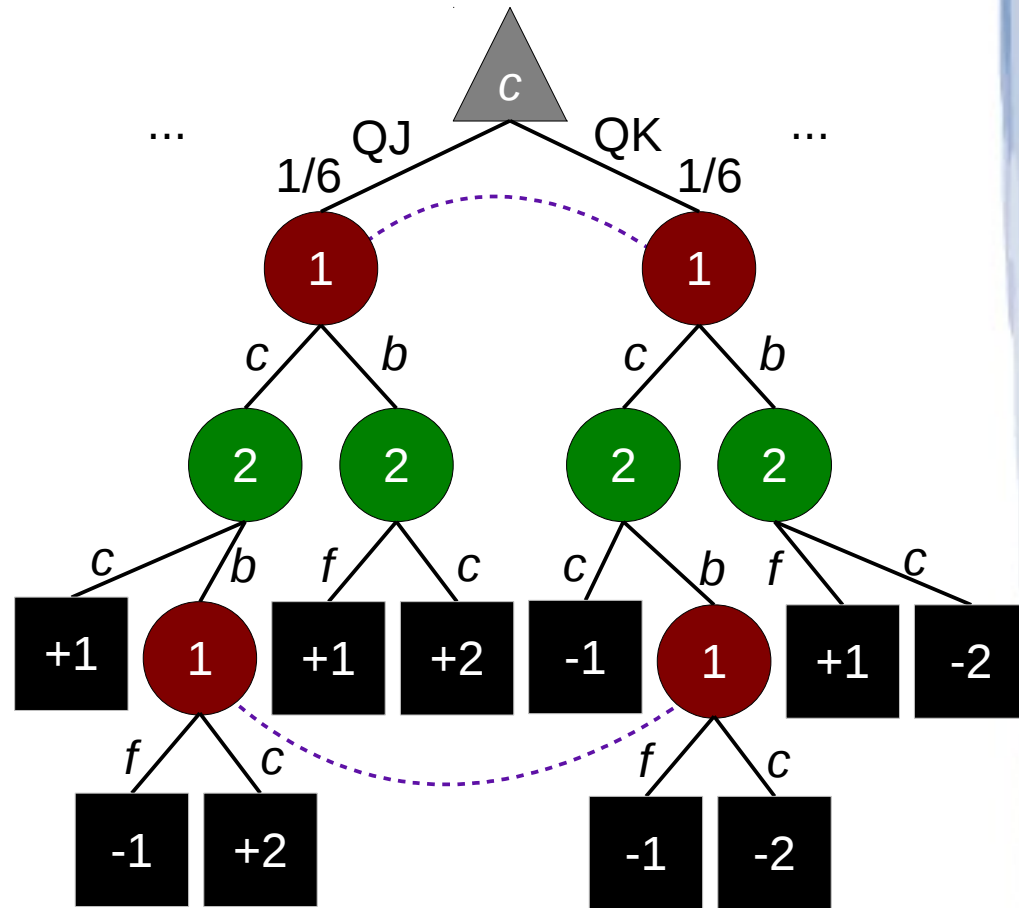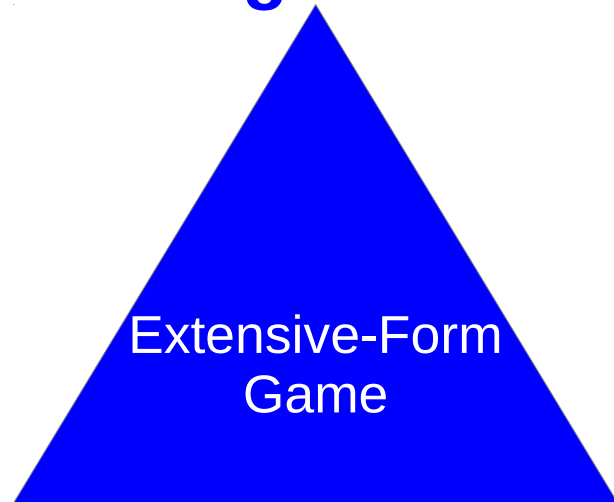
Extensive-Form Game

Background

In general:

Extensive-Form Game

Strategy Profile

Background

In general:

Extensive-Form Game
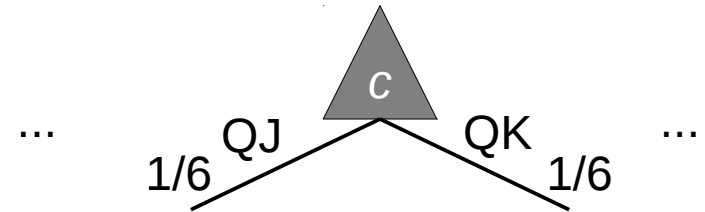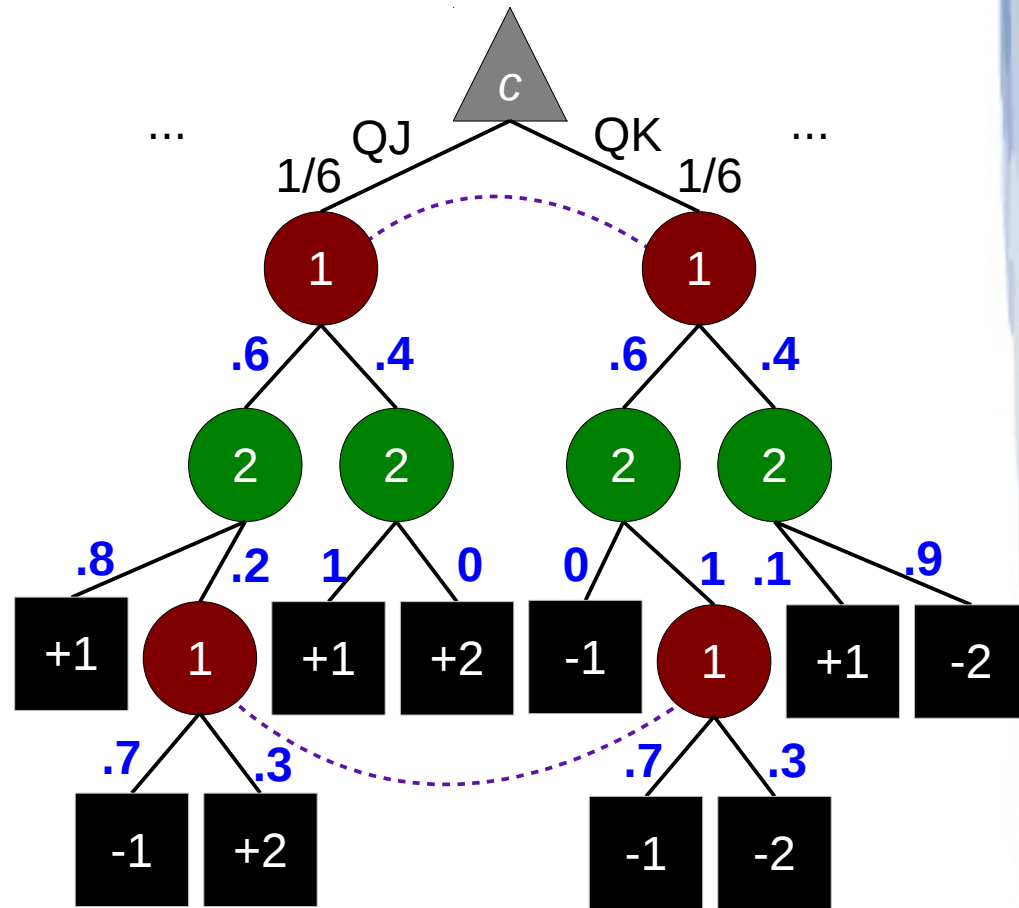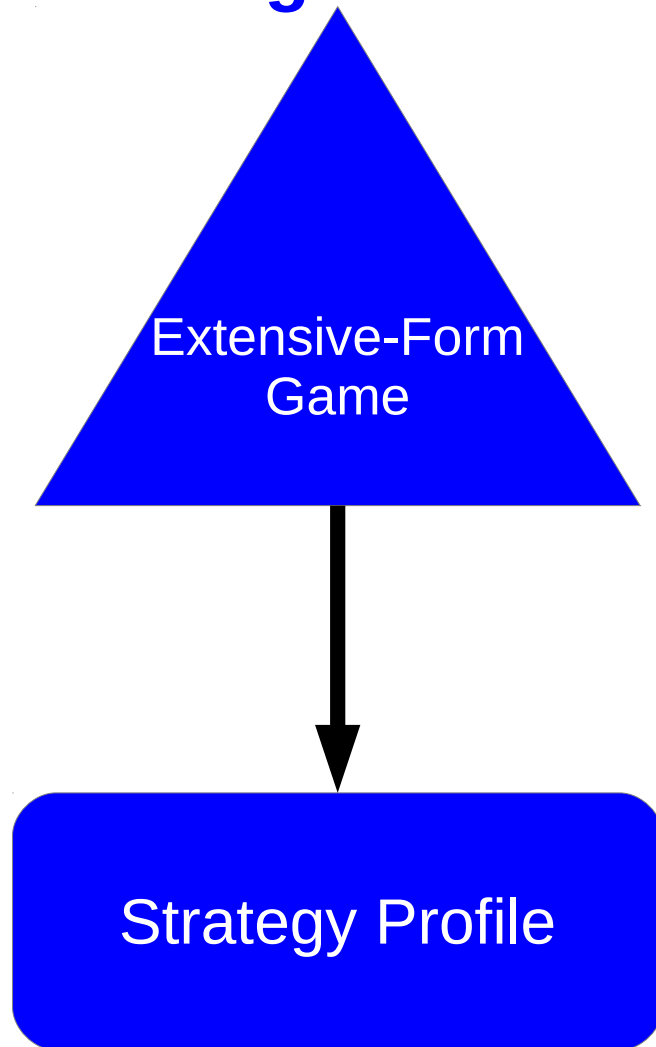
Nash Equilibrium Strategy Profile
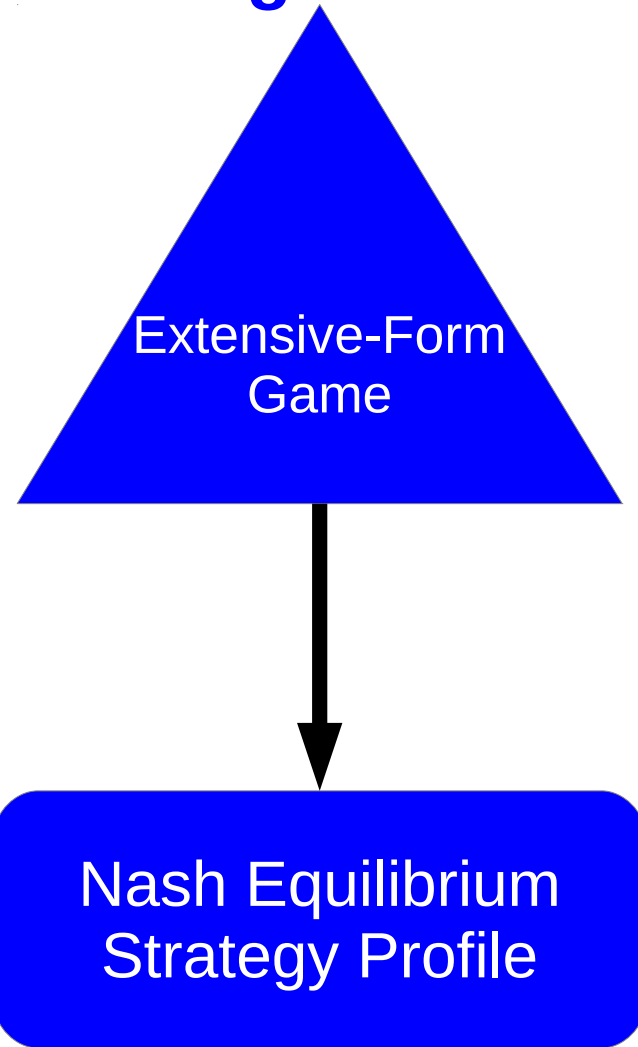
Nash equilibrium:

"No one can change their strategy and do any better."

Background

In general:
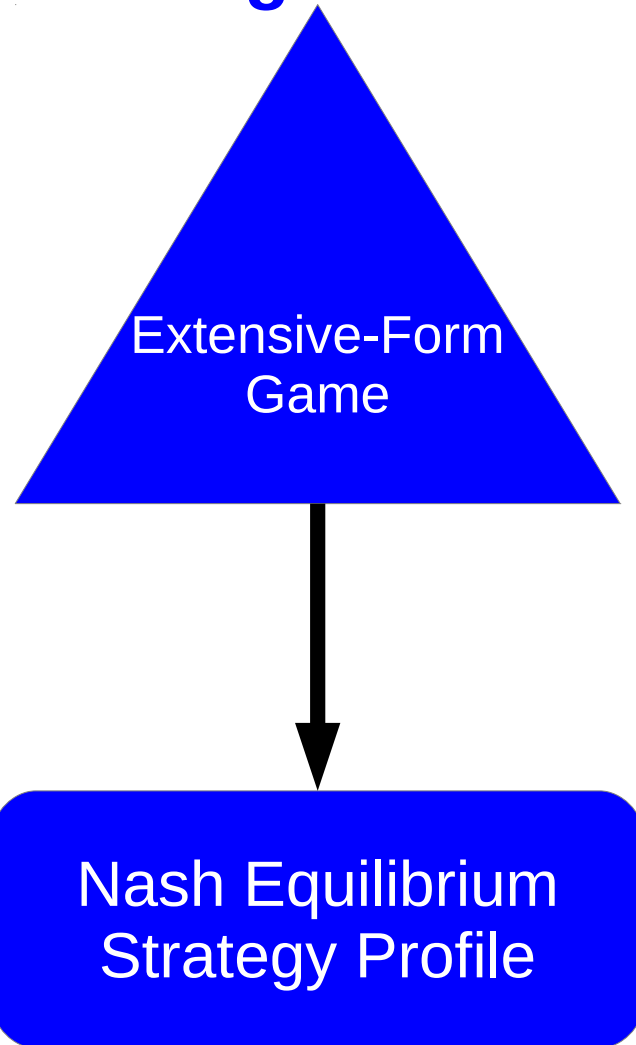
Extensive-Form Game

Nash Equilibrium Strategy Profile

Nash equilibrium:

"No one can change their strategy and do any better."

1/3

1/3

1/3
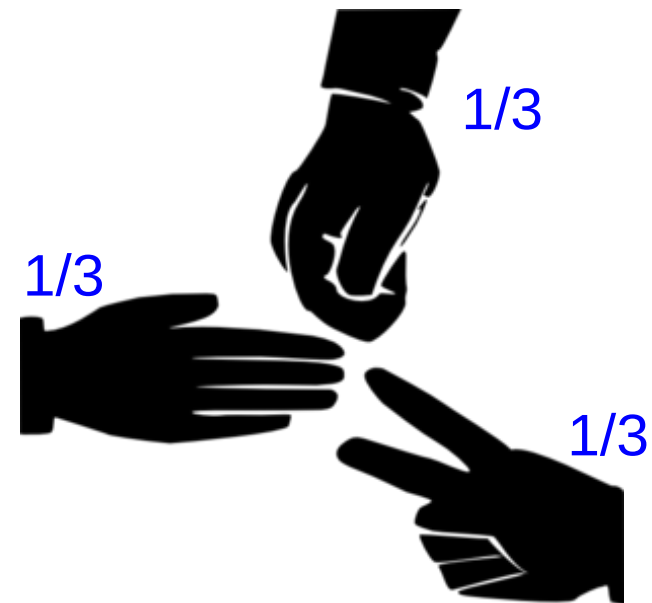
Every game has a Nash equilibrium.

Background

In general:
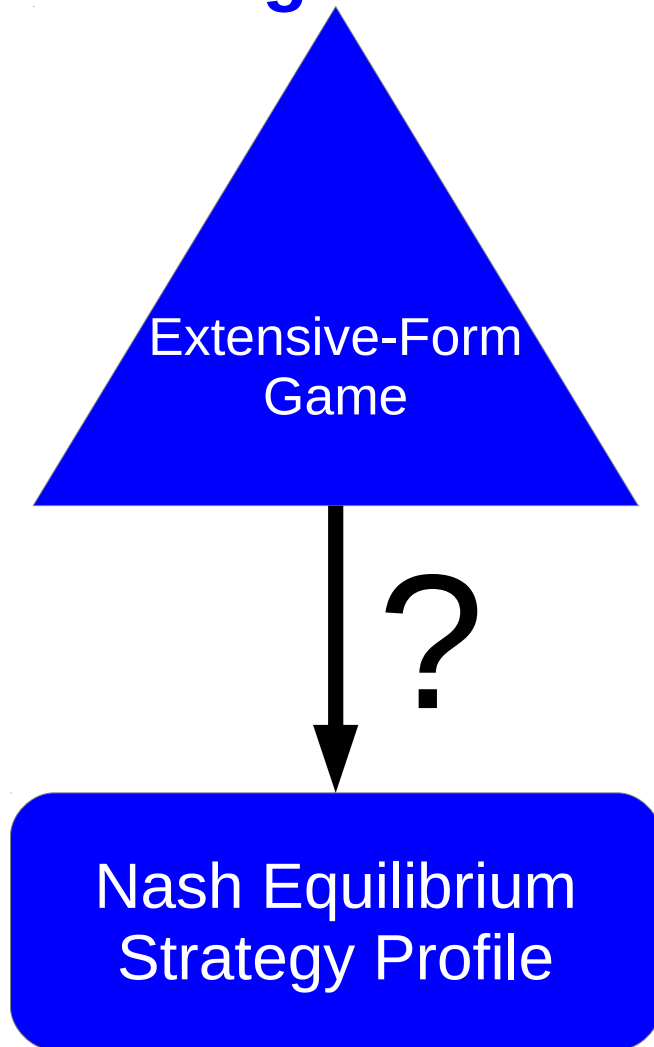
Extensive-Form Game

?

Nash Equilibrium Strategy Profile

Nash equilibrium:

"No one can change their strategy and do any better."

1/3

1/3

1/3
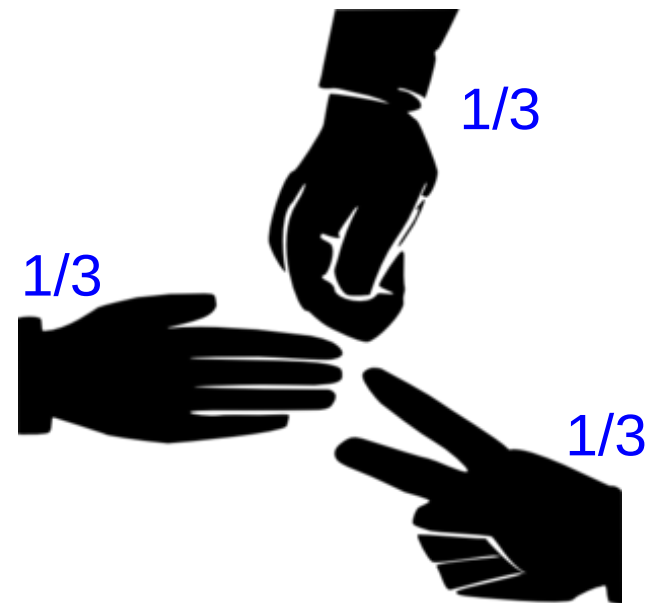
Every game has a Nash equilibrium.

Outline of Presentation

- **Background**
    - **Counterfactual Regret Minimization (CFR)**
- Theoretical Advancements for CFR in:
    - Many player games
    - Imperfect recall games
- CFR Speed-Ups
- Tricks for CFR with Memory Limitations
- Conclusion + Future Work

CFR

- "The alpha-beta search of imperfect information games."

CFR
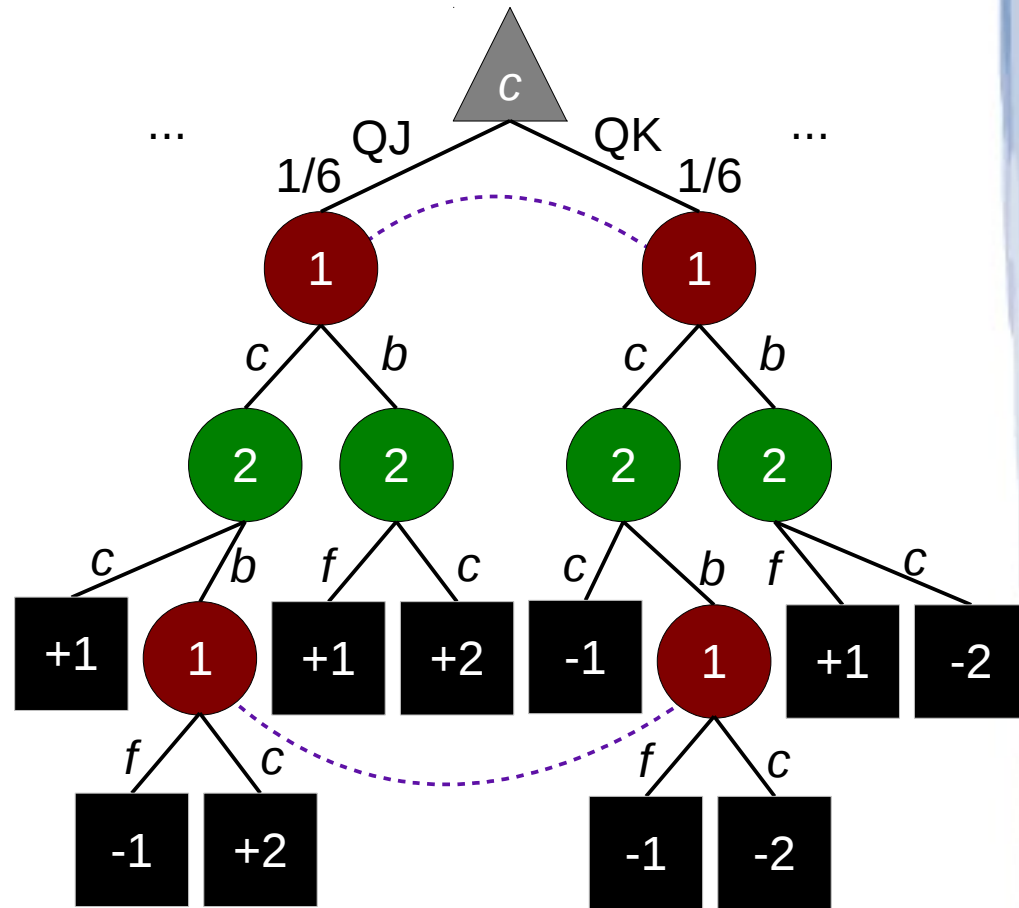
- "The alpha-beta search of imperfect information games."
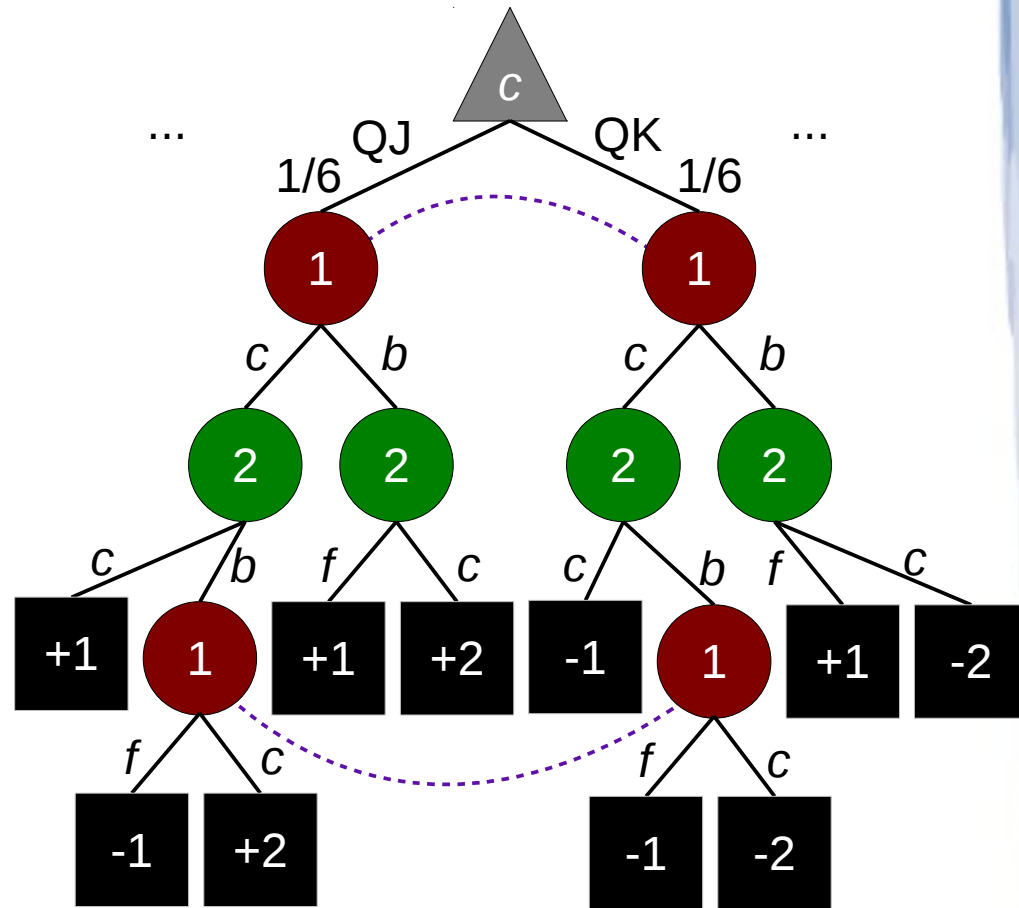
- Offline algorithm

CFR

- "The alpha-beta search of imperfect information games."

- Offline algorithm

- Iterative, "self-play"

CFR
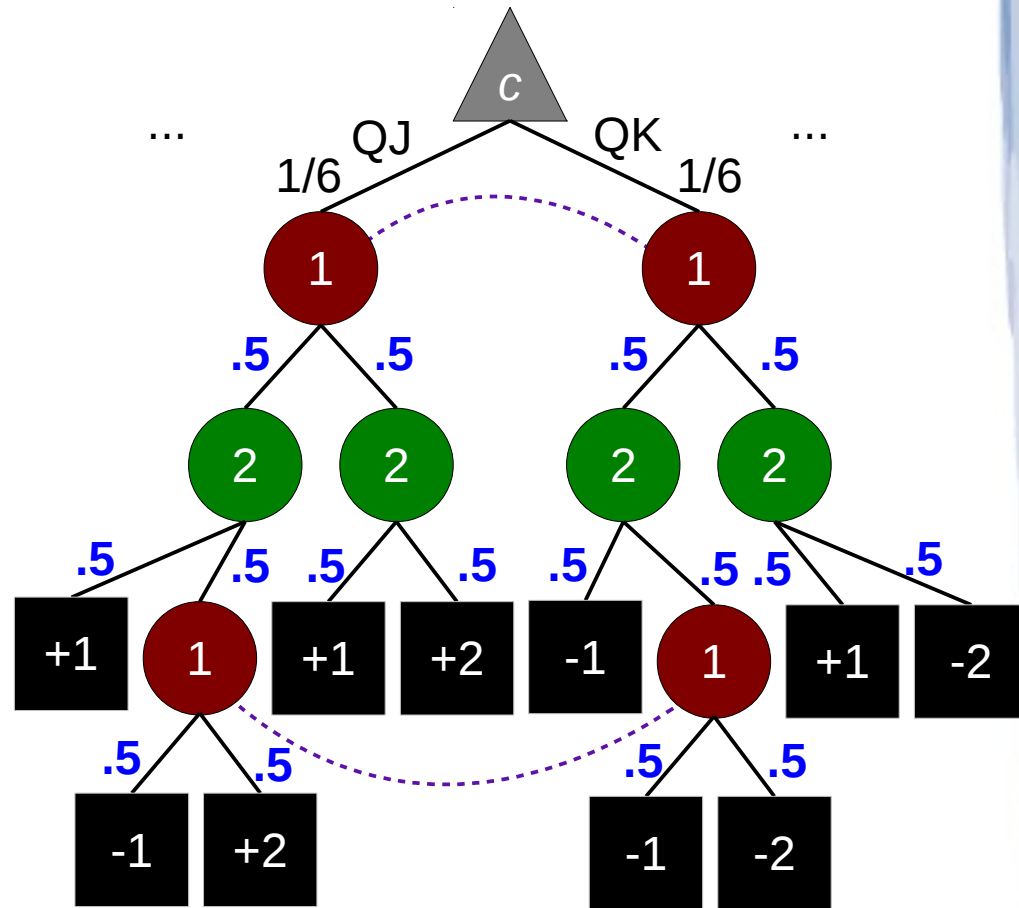
- "The alpha-beta search of imperfect information games."

- Offline algorithm

- Iterative, "self-play"

- For each iteration, update action probabilities at every information set.

CFR

$$\frac{\text{Strategy 1} + \text{Strategy 2} + ... + \text{Strategy } T}{T}$$

$T \longrightarrow \infty$

Nash Equilibrium Strategy Profile

$=$

Average Strategy Profile

$T = $ number of iterations

Background

Background

Background

Texas Hold'em

$>10^{14}$ information sets

CFR

Nash Equilibrium Strategy Profile

> 5 million GB

Background

Large
Extensive-Form
Game

?

Nash Equilibrium
Strategy Profile
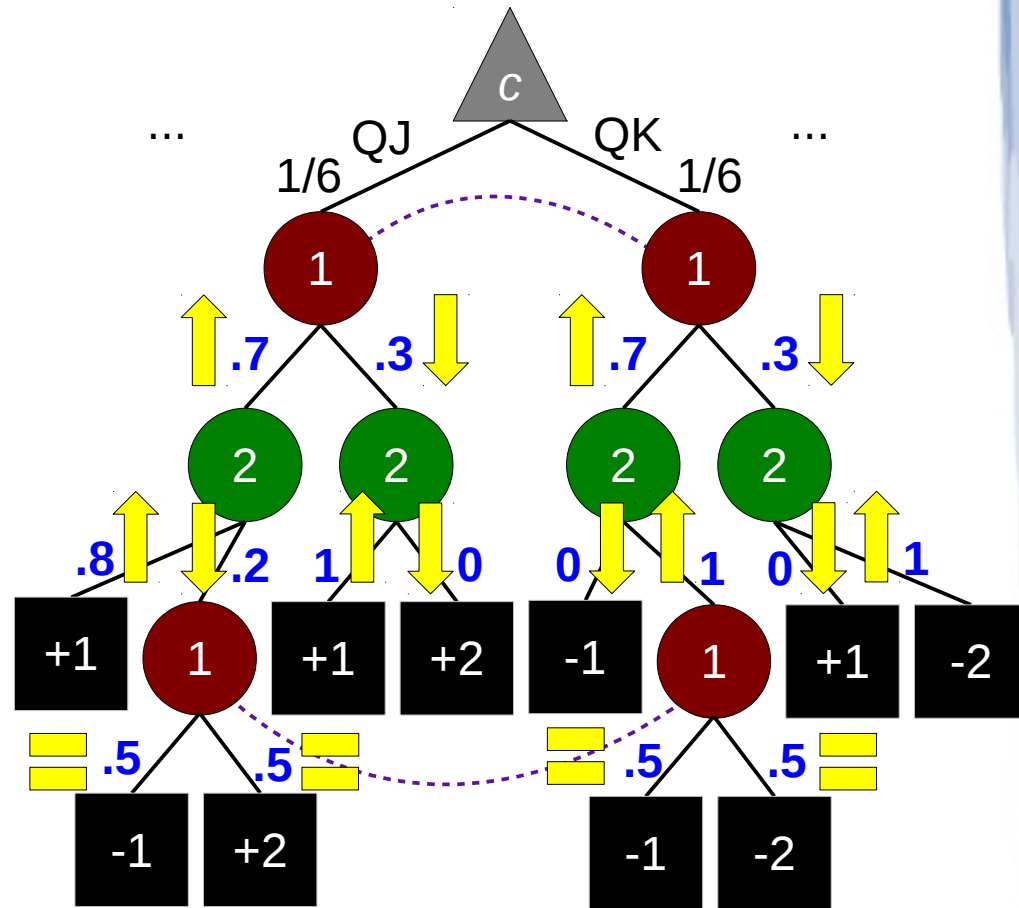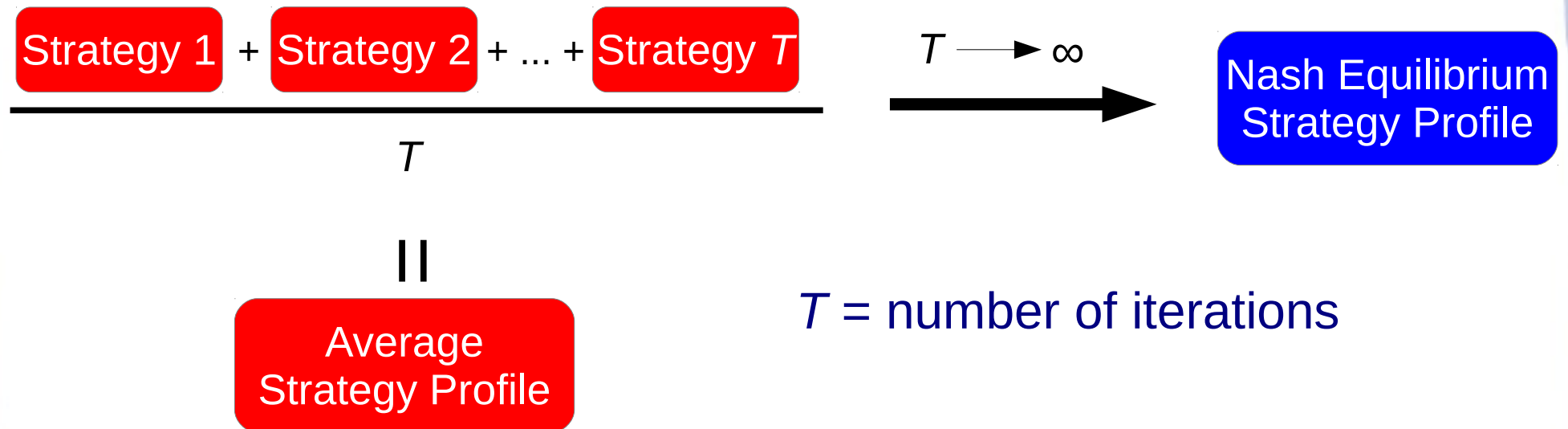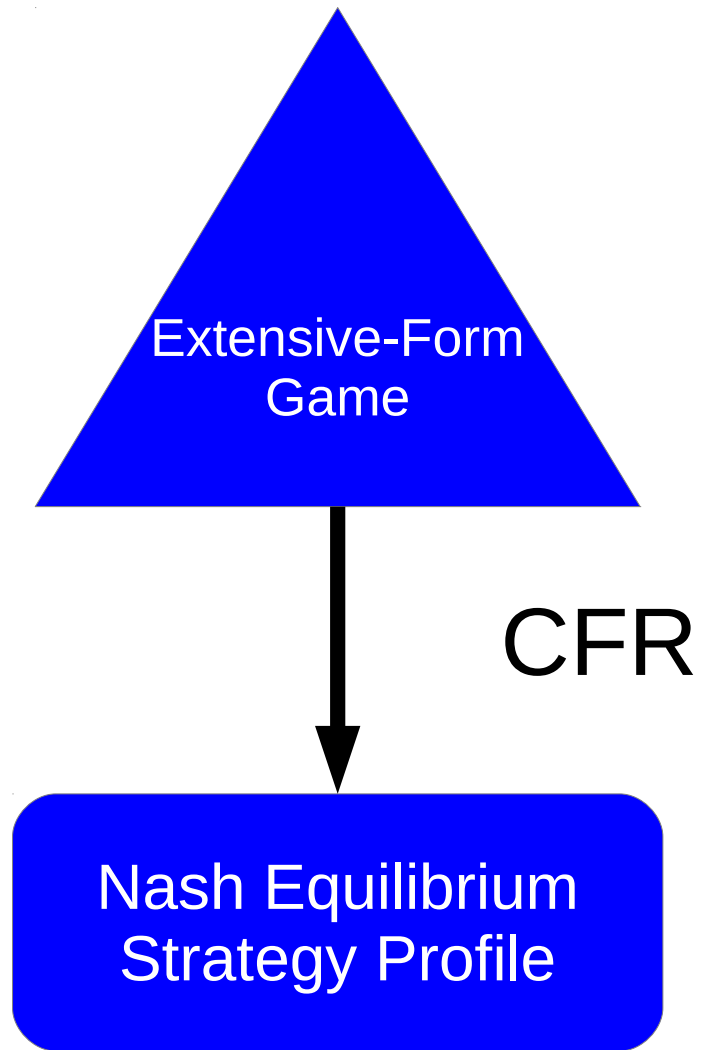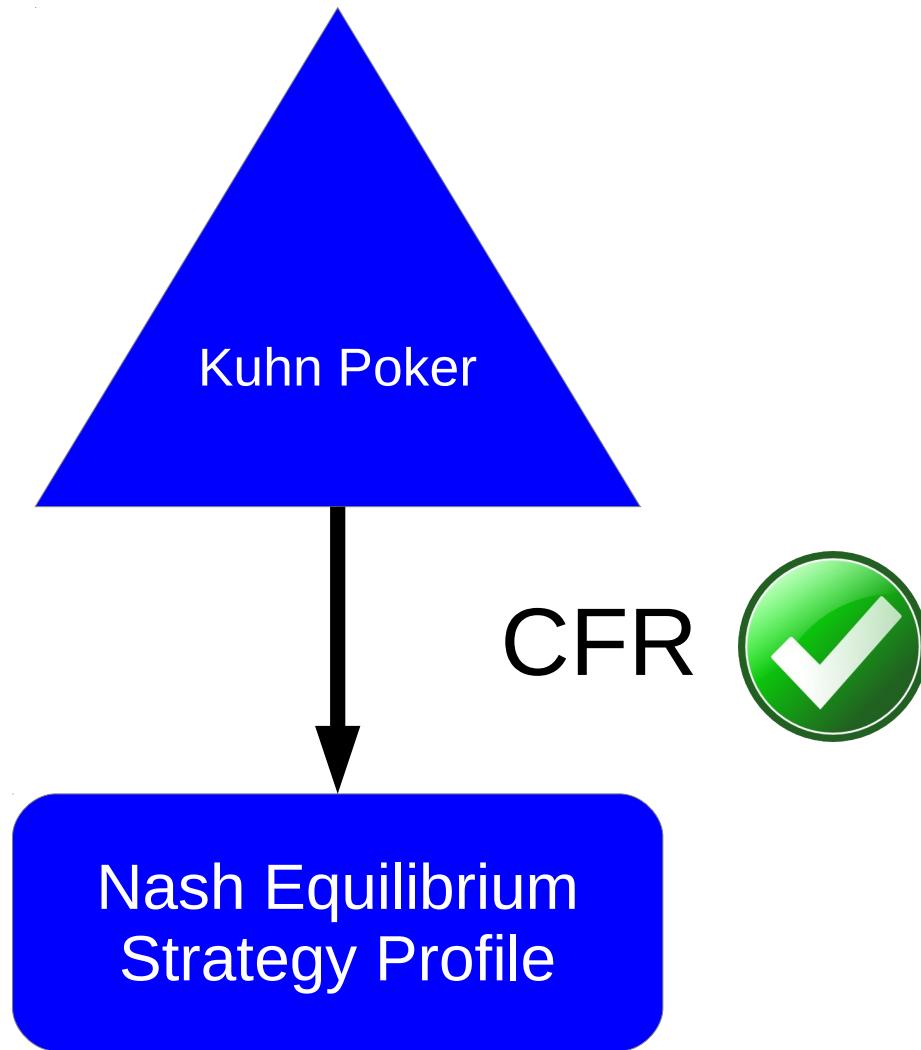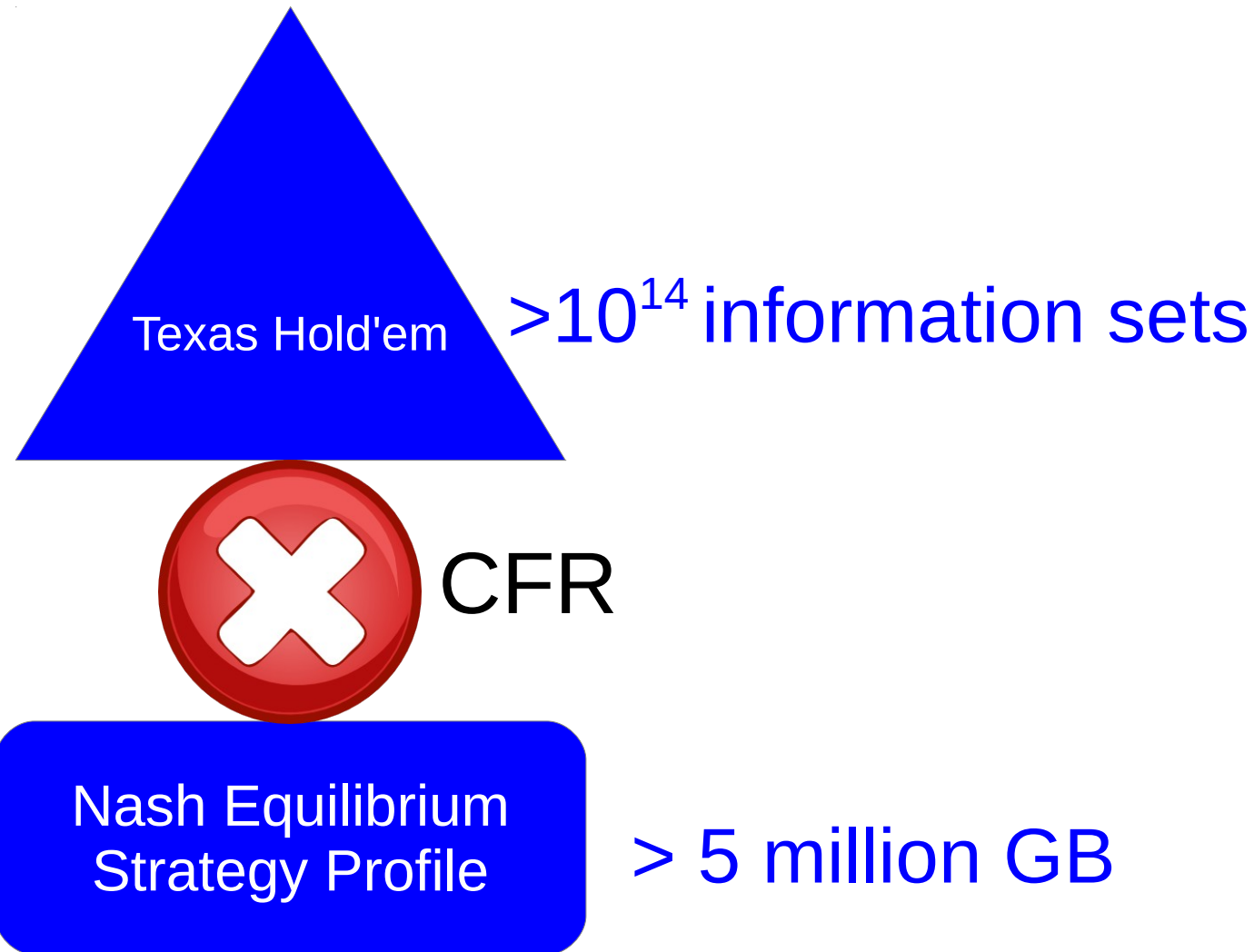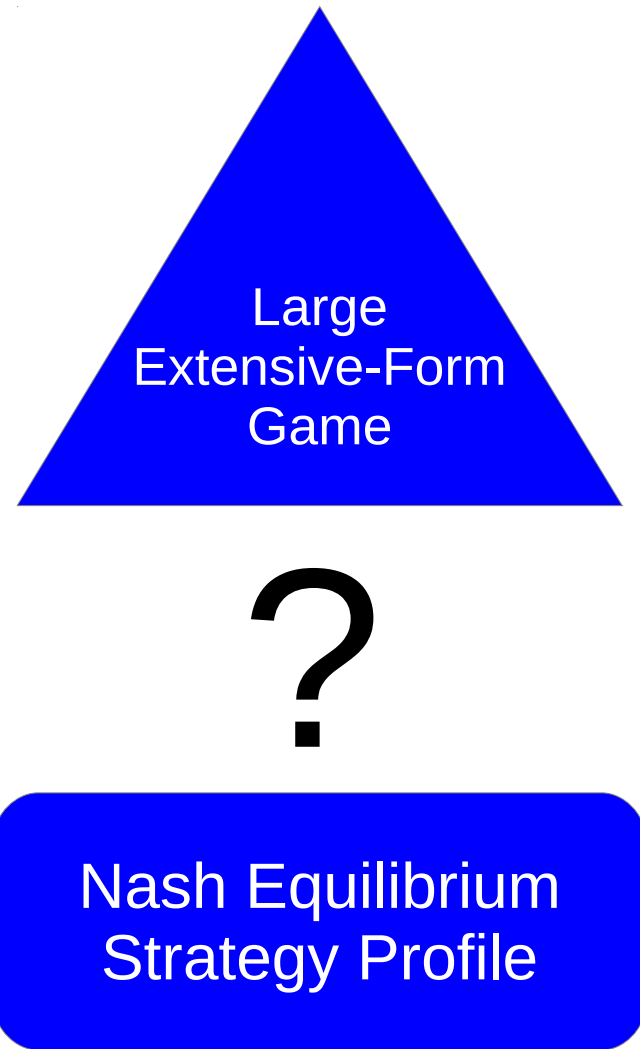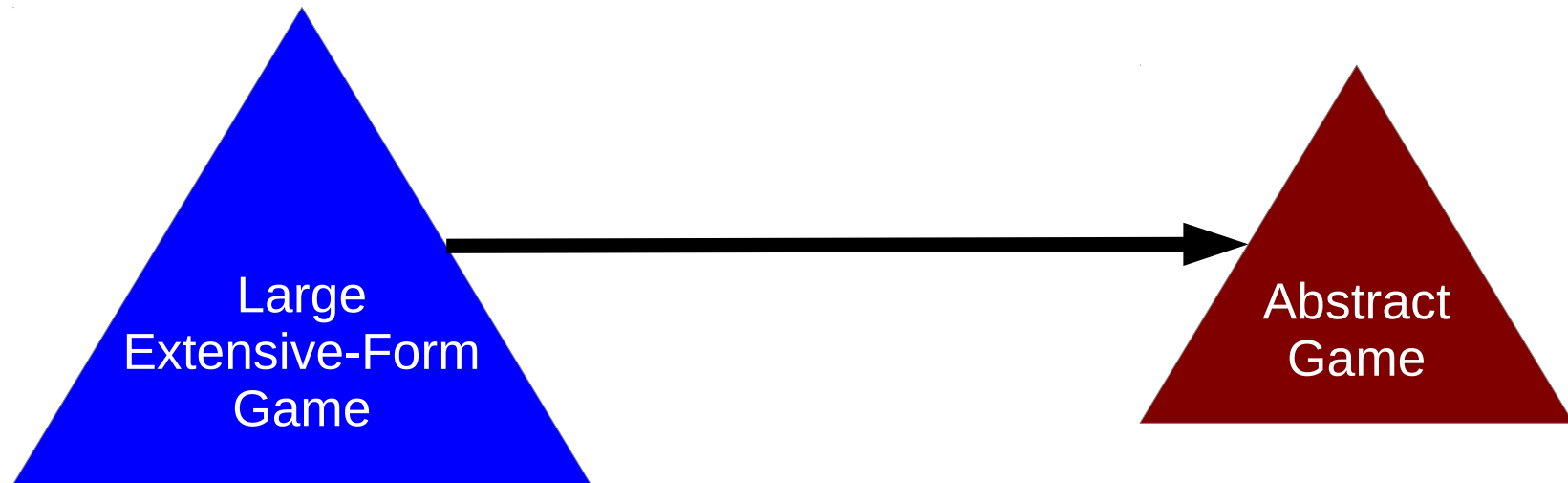
Background

Background

- Merge card deals into buckets.
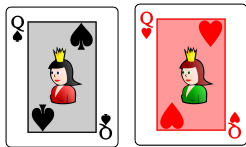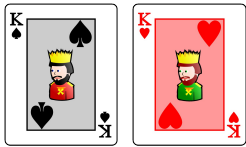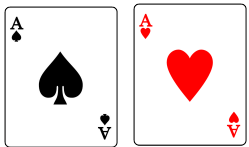
Extensive-Form Game
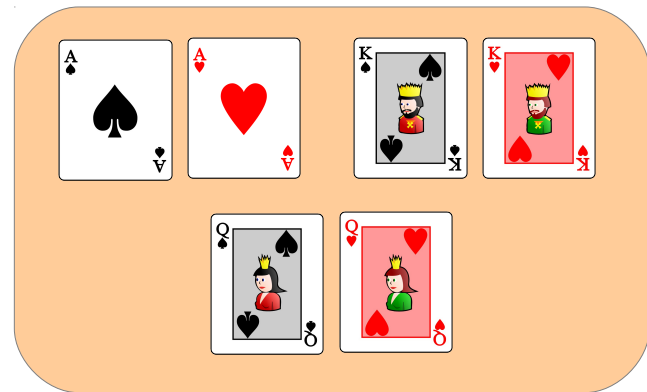
Abstract Game

Background

- Merge card deals into buckets.
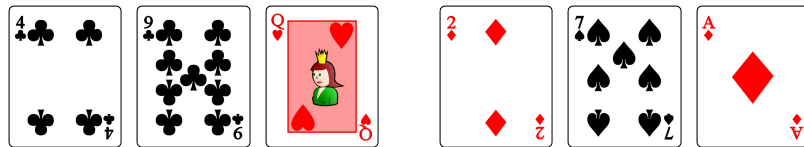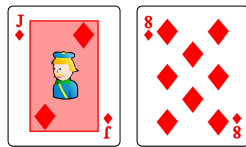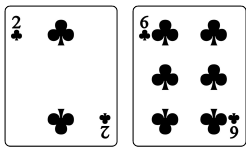
Extensive-Form Game
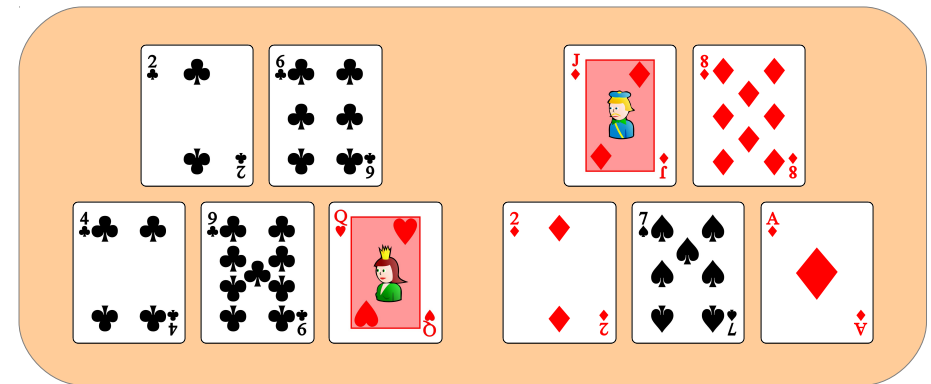
Abstract Game
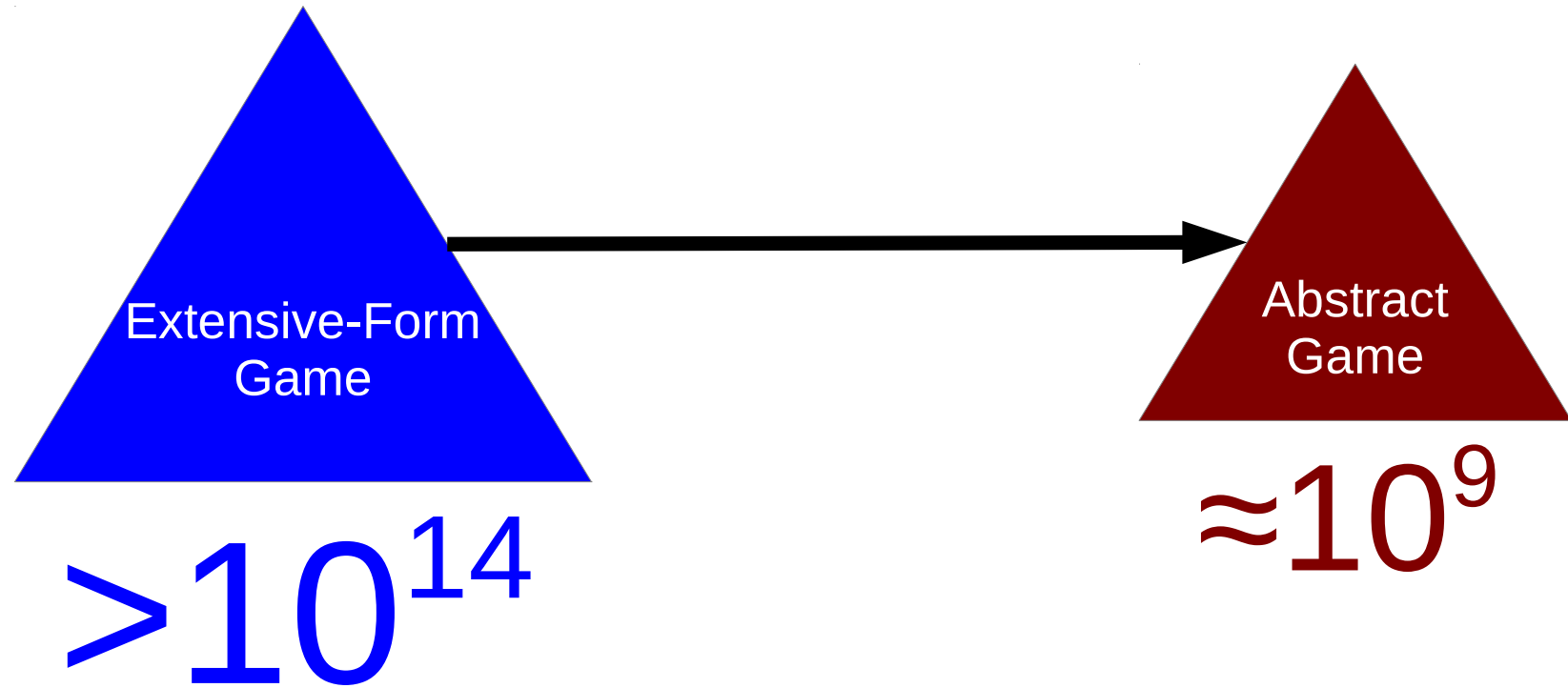
Background

Background
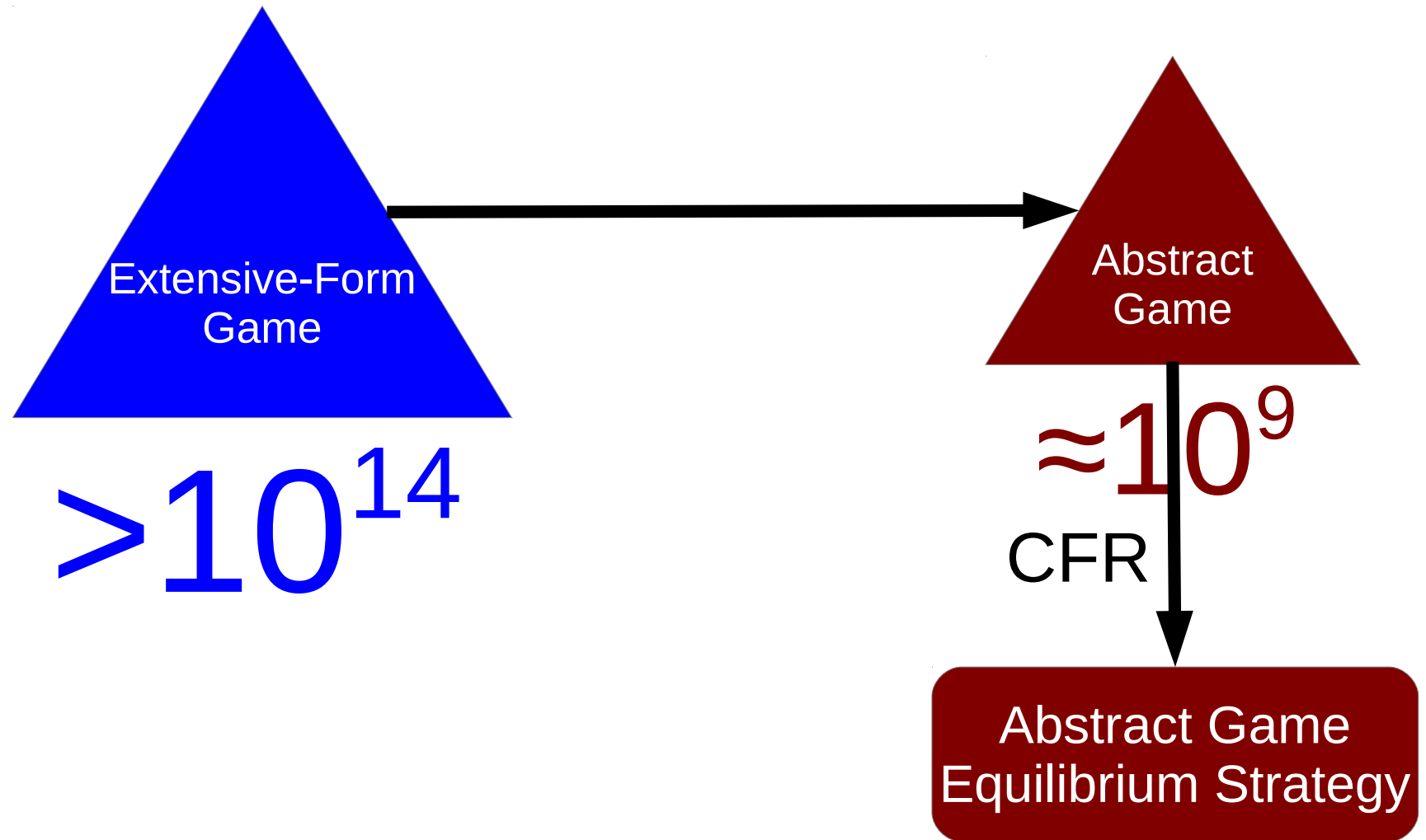
Background

Outline of Presentation

- Background

  – Counterfactual Regret Minimization (CFR)

- **Theoretical Advancements for CFR in:**

  – **Many player games**

  – Imperfect recall games

- CFR Speed-Ups

- Tricks with Memory Limitations

- Conclusion + Future Work

Theory – Many Player Games

Extensive-Form
Game

CFR

Nash Equilibrium
Strategy Profile
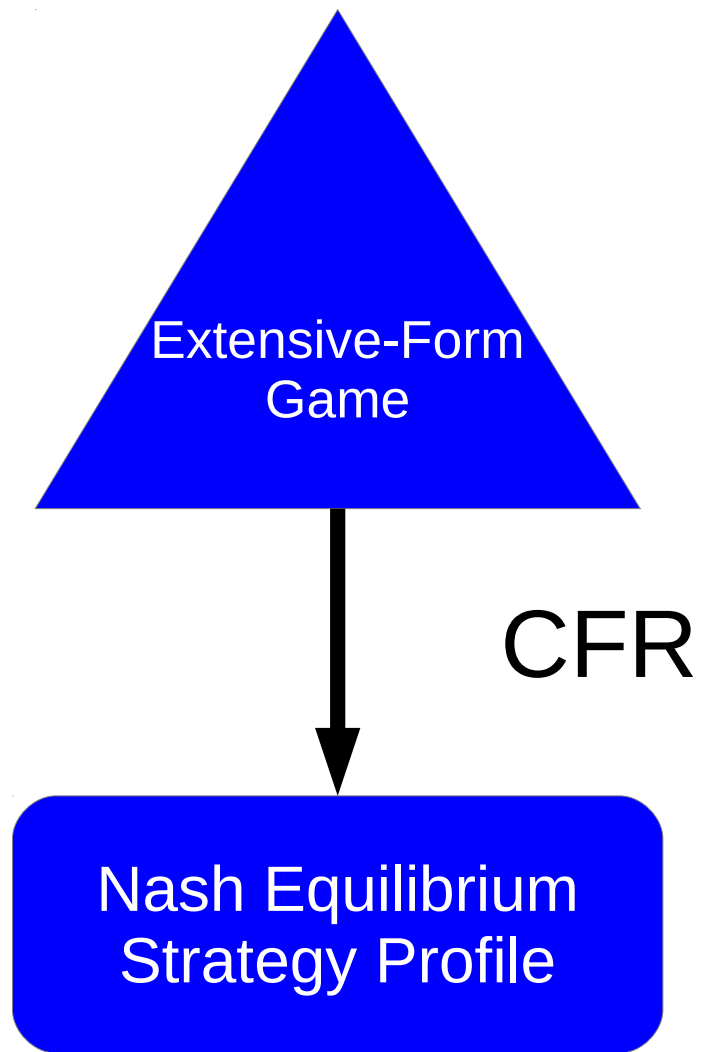
LIAR!

Theory – Many Player Games
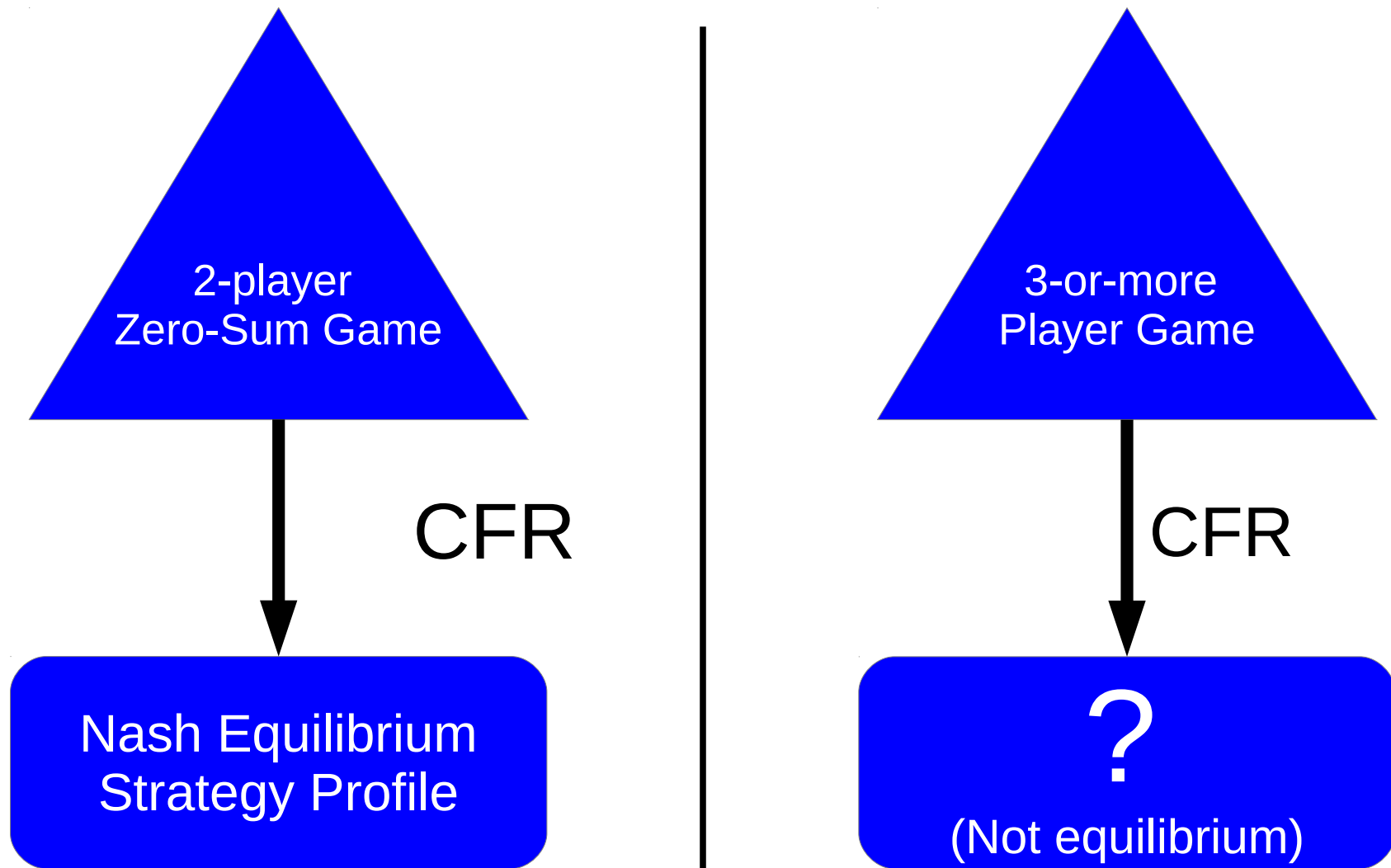
Theory – Many Player Games

Annual Computer Poker Competition
3-Player Limit Texas Hold'em - 2009

Agent	Total Bankroll (mbb/g)
Hyperborean3p	319 ± 2
dpp	171 ± 2
akuma	151 ± 2
CMURingLimit	-37 ± 2
dcu3pl	-63 ± 2
Bluechip	-548 ± 2
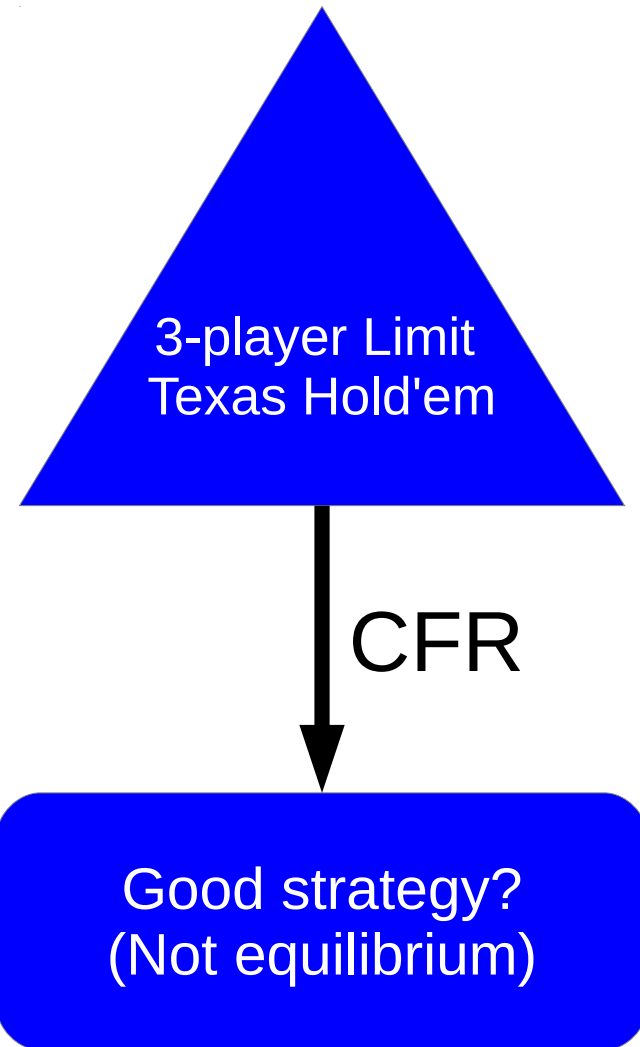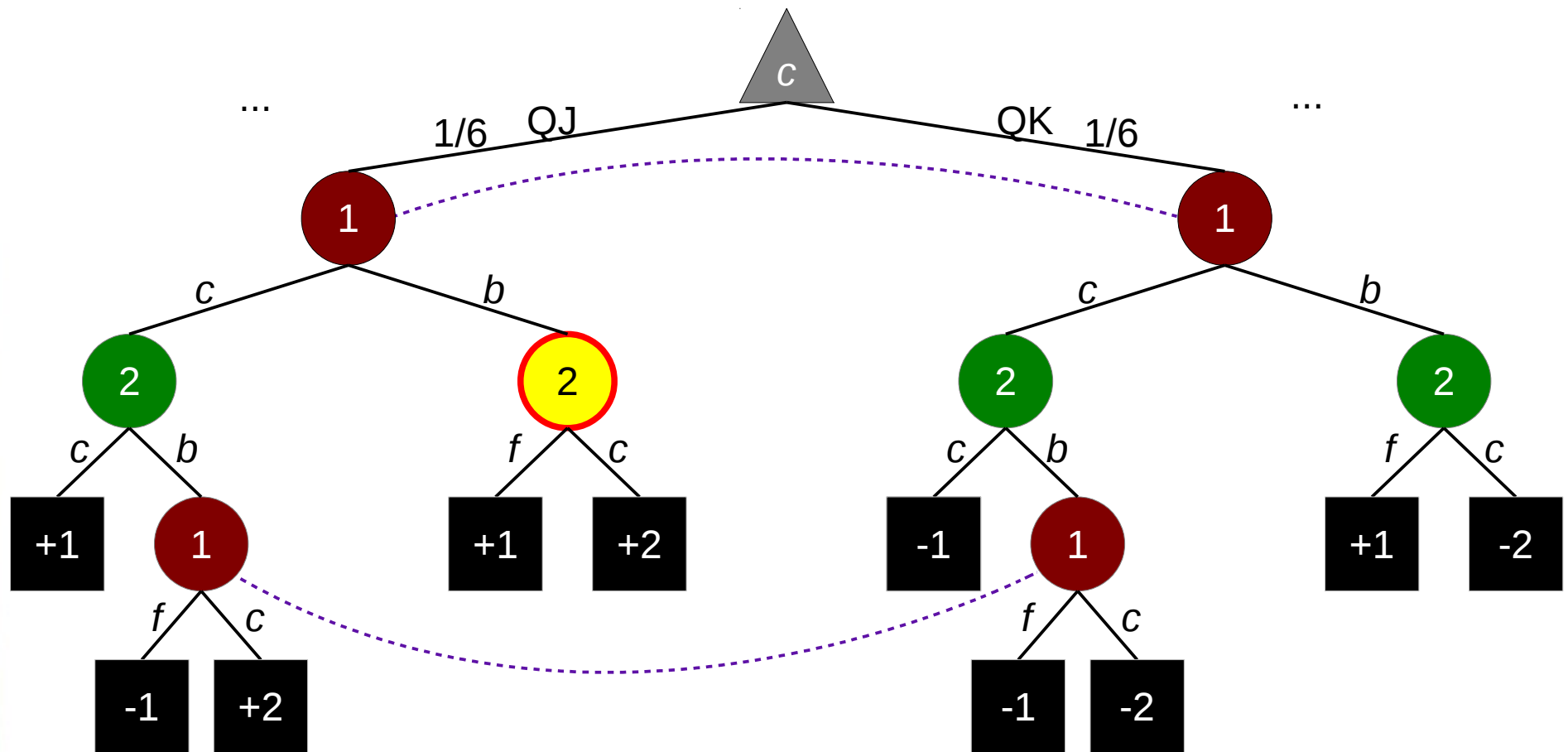
3-player Limit
Texas Hold'em

CFR

Good strategy?
(Not equilibrium)

Theory – Many Player Games

Theory – Many Player Games

Theory – Many Player Games
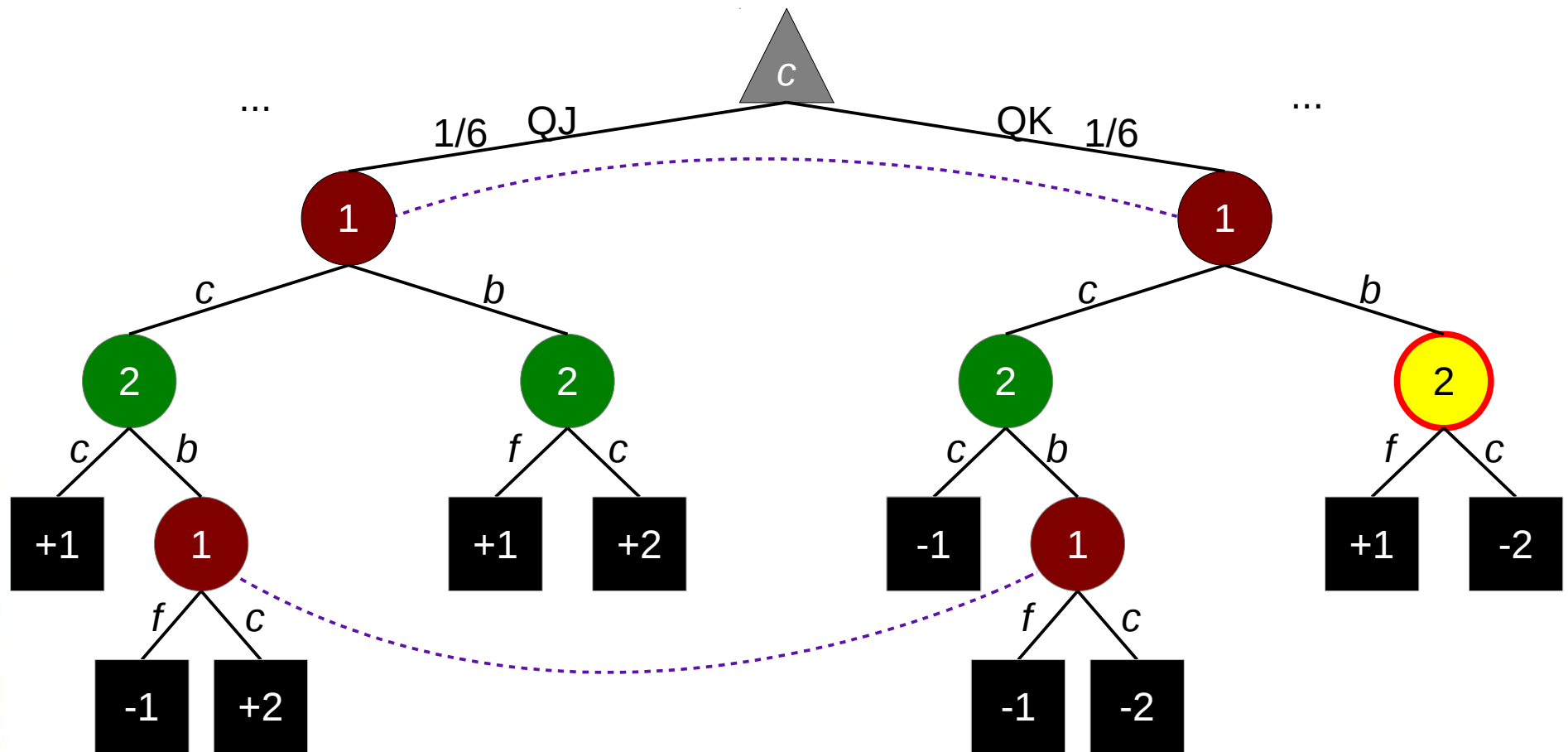
Dominated Strategies
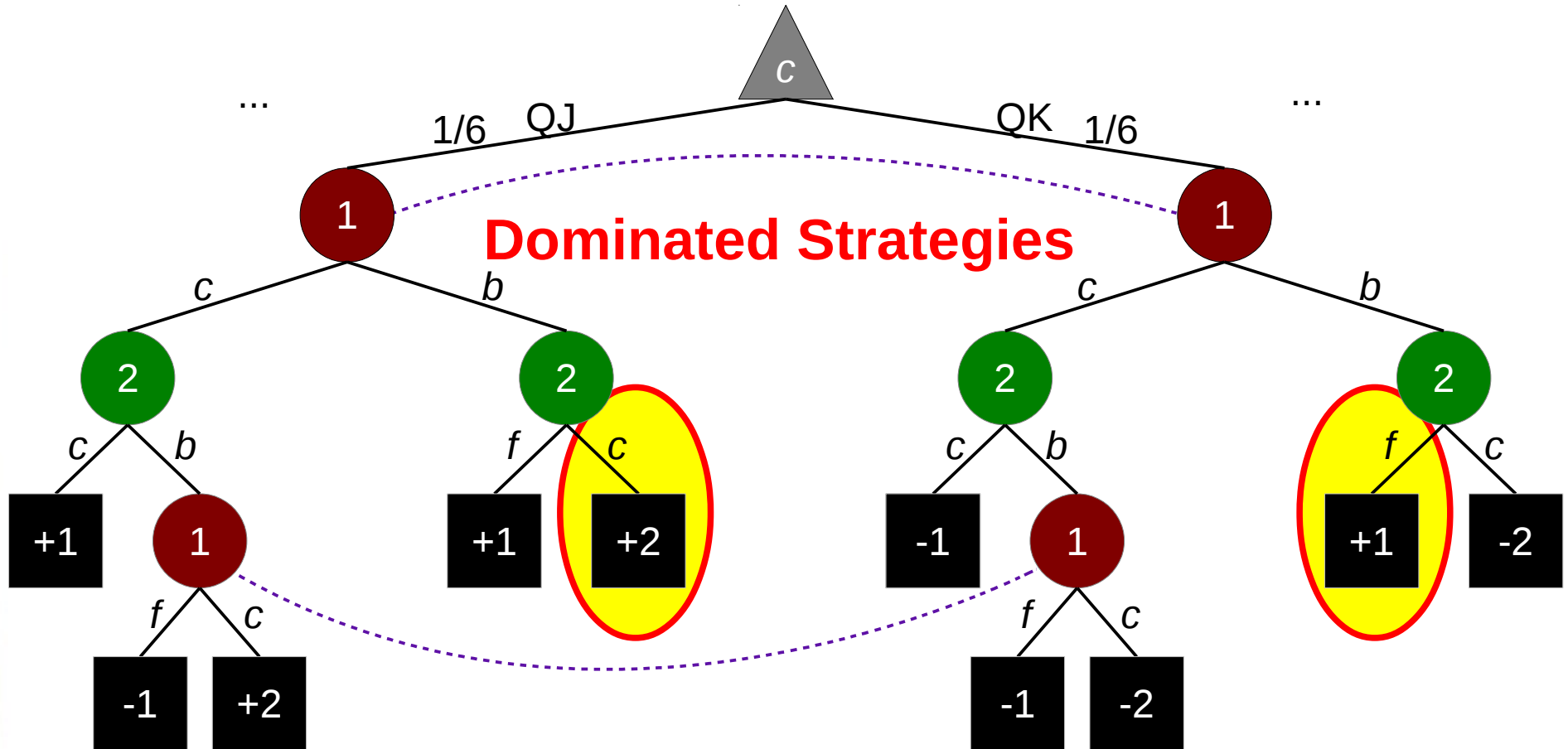
Theory – Many Player Games
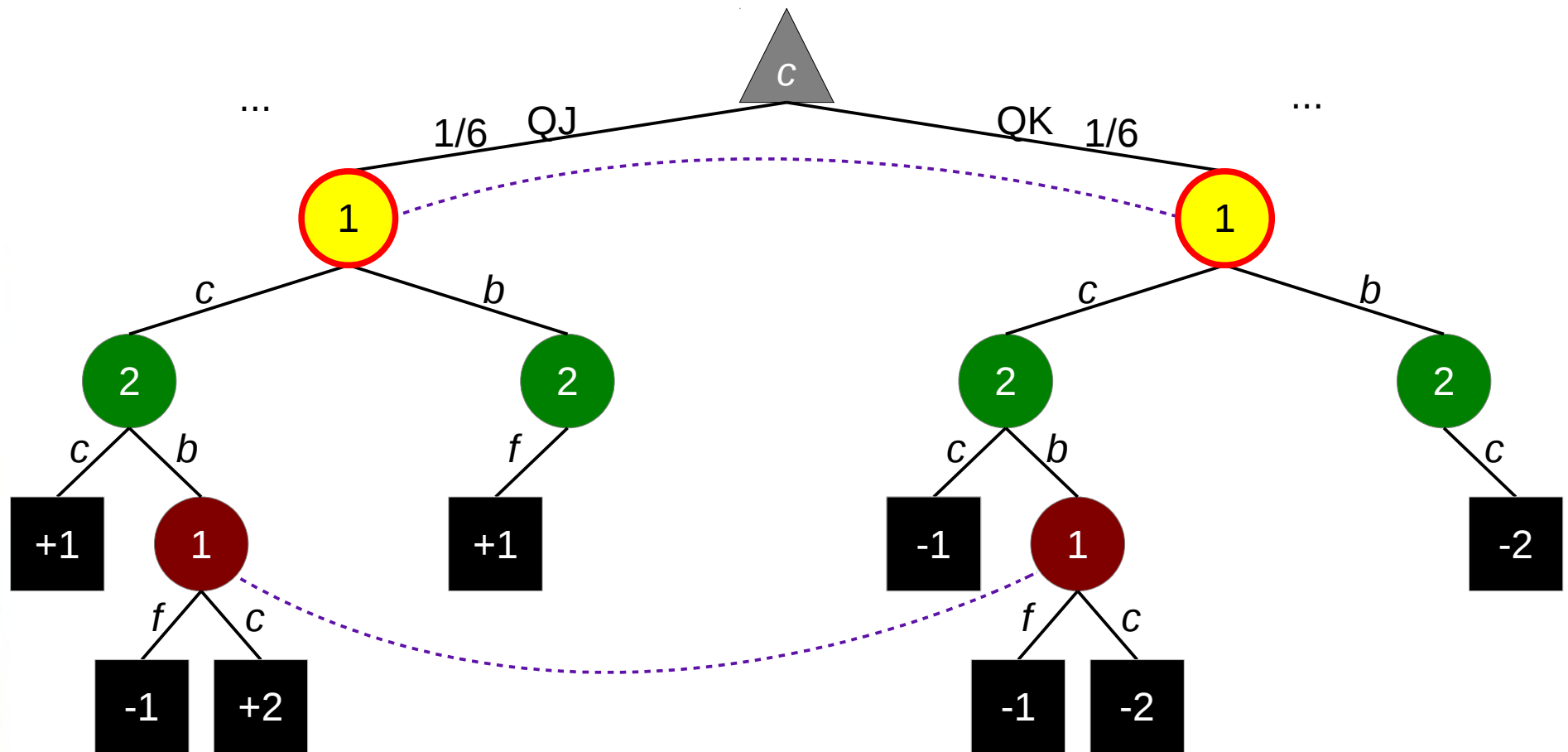
Theory – Many Player Games
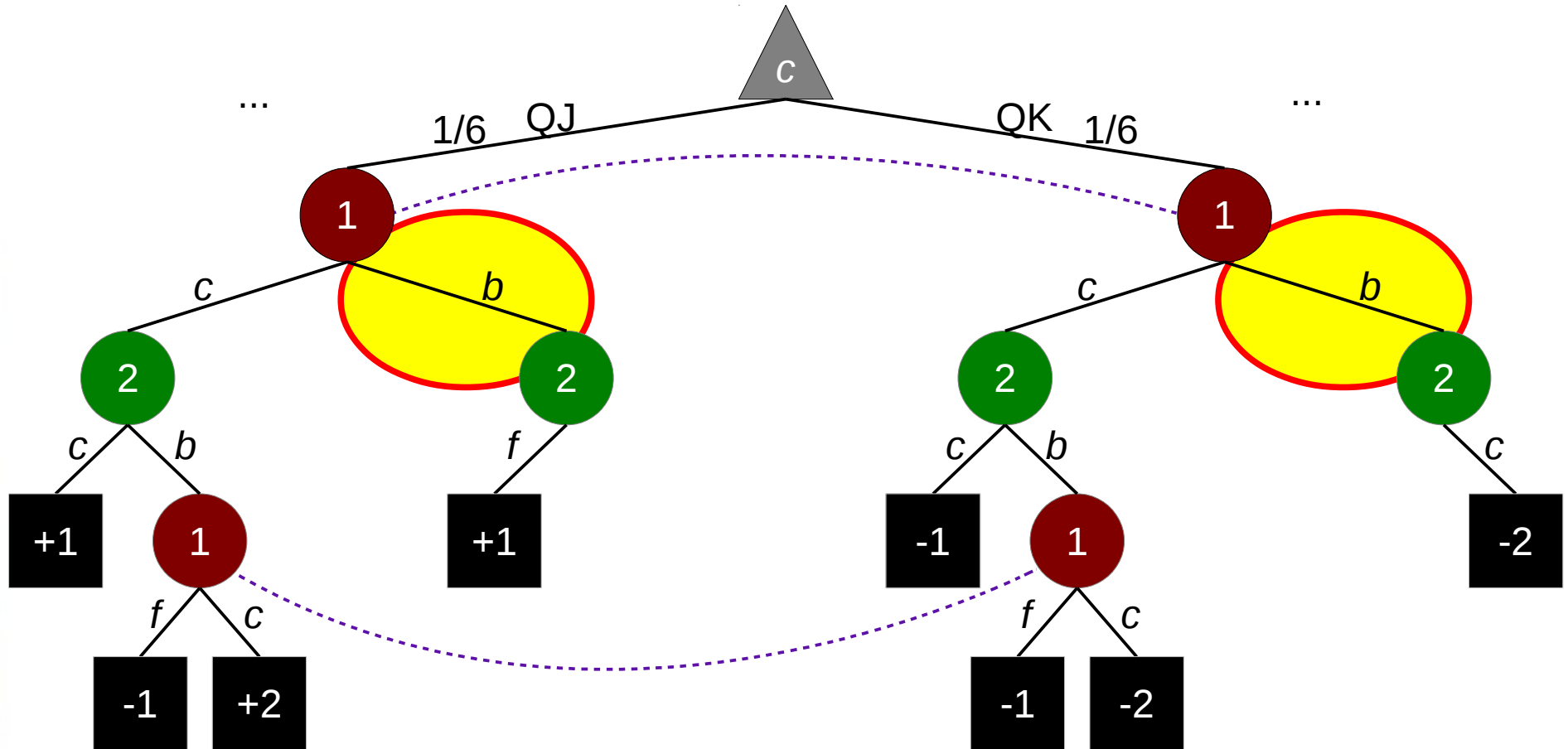
Iteratively Dominated Strategy

Theory – Many Player Games
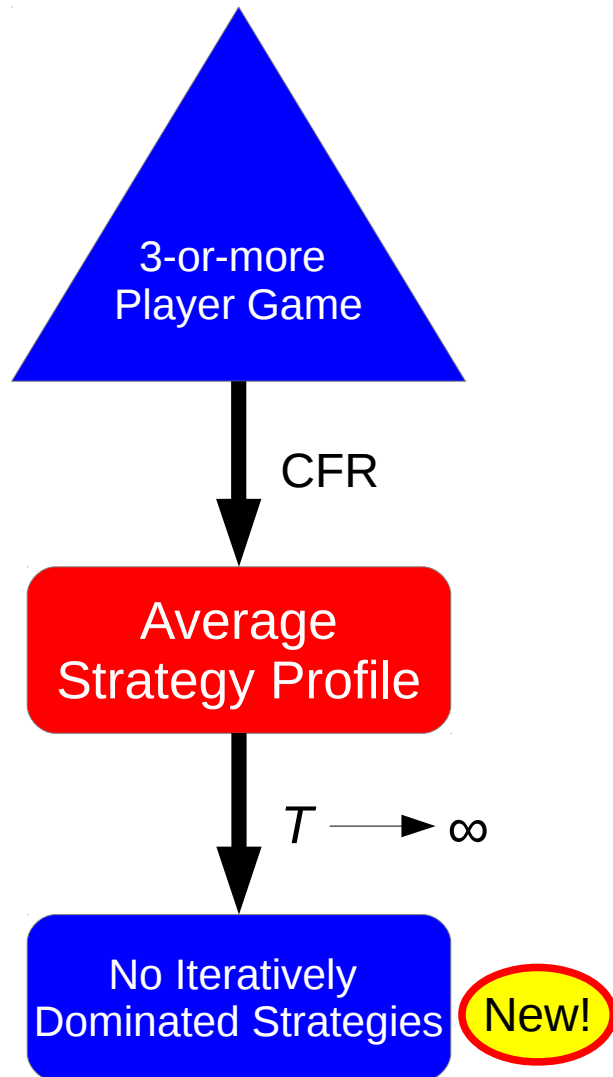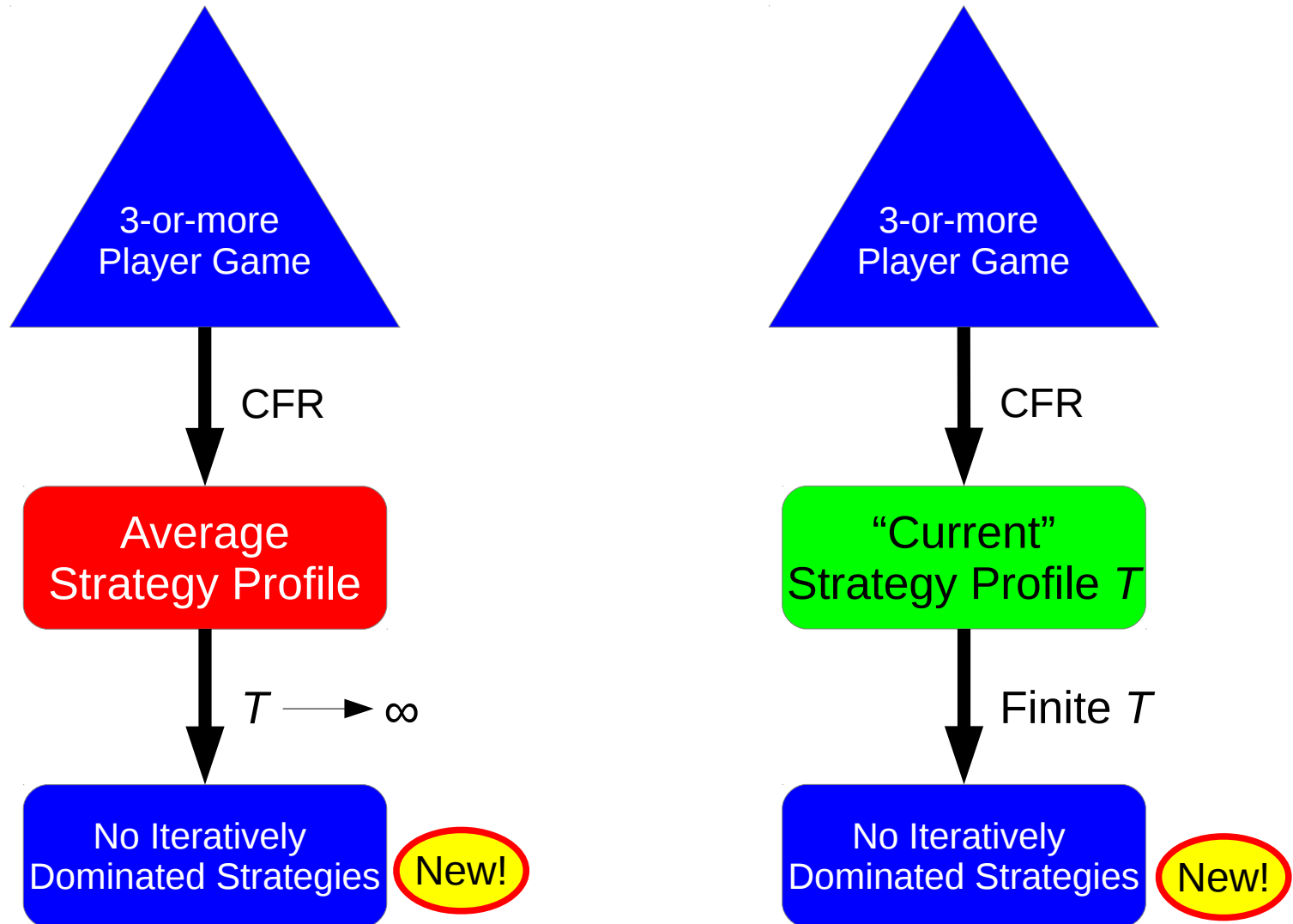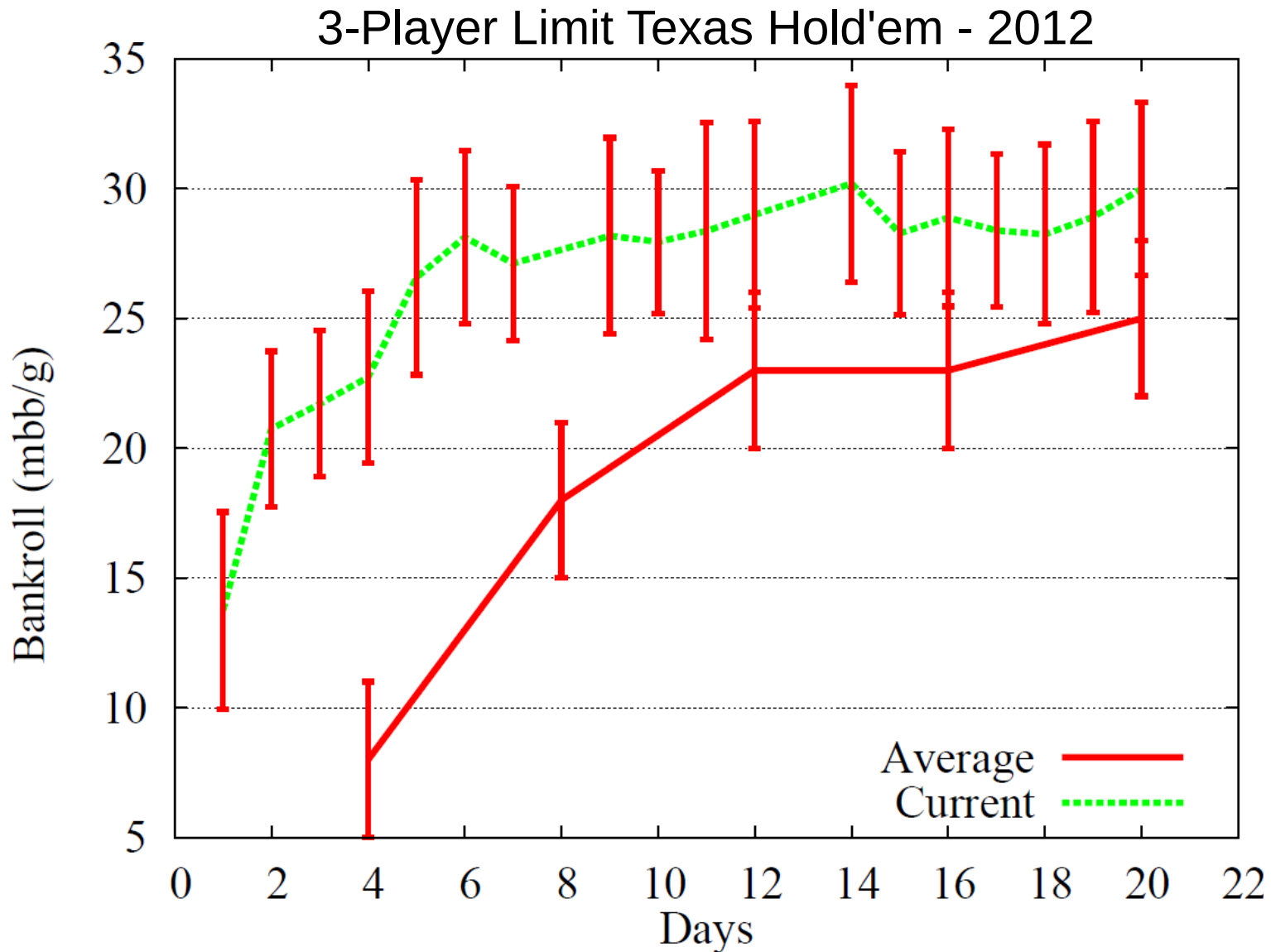
3-or-more Player Game

CFR

Average Strategy Profile

$T \longrightarrow \infty$

No Iteratively Dominated Strategies   New!

[G., arXiv ePrints 2013]

Theory – Many Player Games

3-or-more Player Game → **CFR** → **Average Strategy Profile** → $T \longrightarrow \infty$ → No Iteratively Dominated Strategies (New!)

3-or-more Player Game → **CFR** → **"Current" Strategy Profile *T*** → Finite *T* → No Iteratively Dominated Strategies (New!)

[G., arXiv ePrints 2013]

Theory – Many Player Games

3-Player Limit Texas Hold'em - 2012

New!

Outline of Presentation

- Background

  - Counterfactual Regret Minimization (CFR)

- **Theoretical Advancements for CFR in:**

  - Many player games

  - **Imperfect recall games**

- CFR Speed-Ups

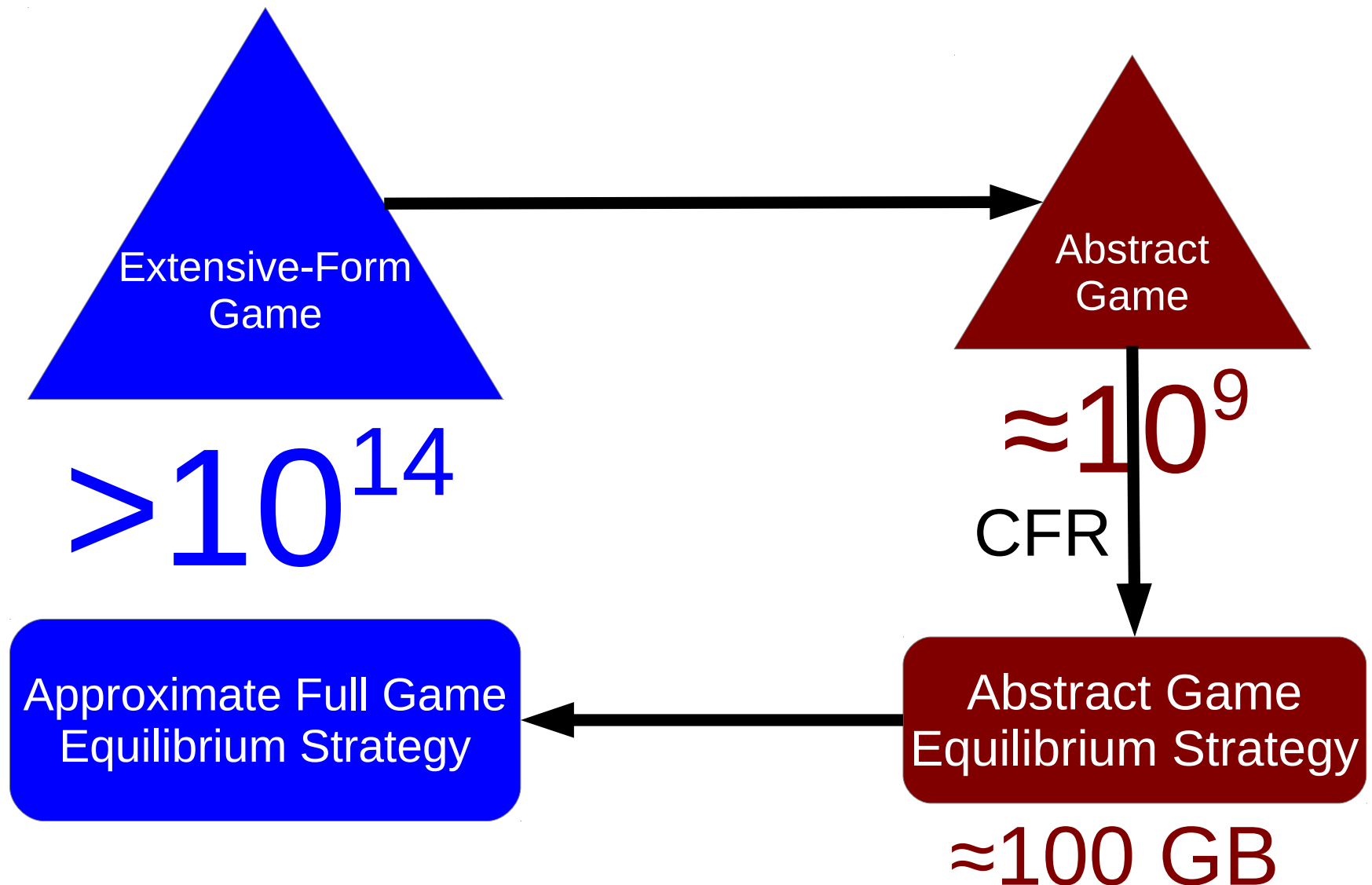- Tricks with Memory Limitations

- Conclusion + Future Work

Imperfect Recall
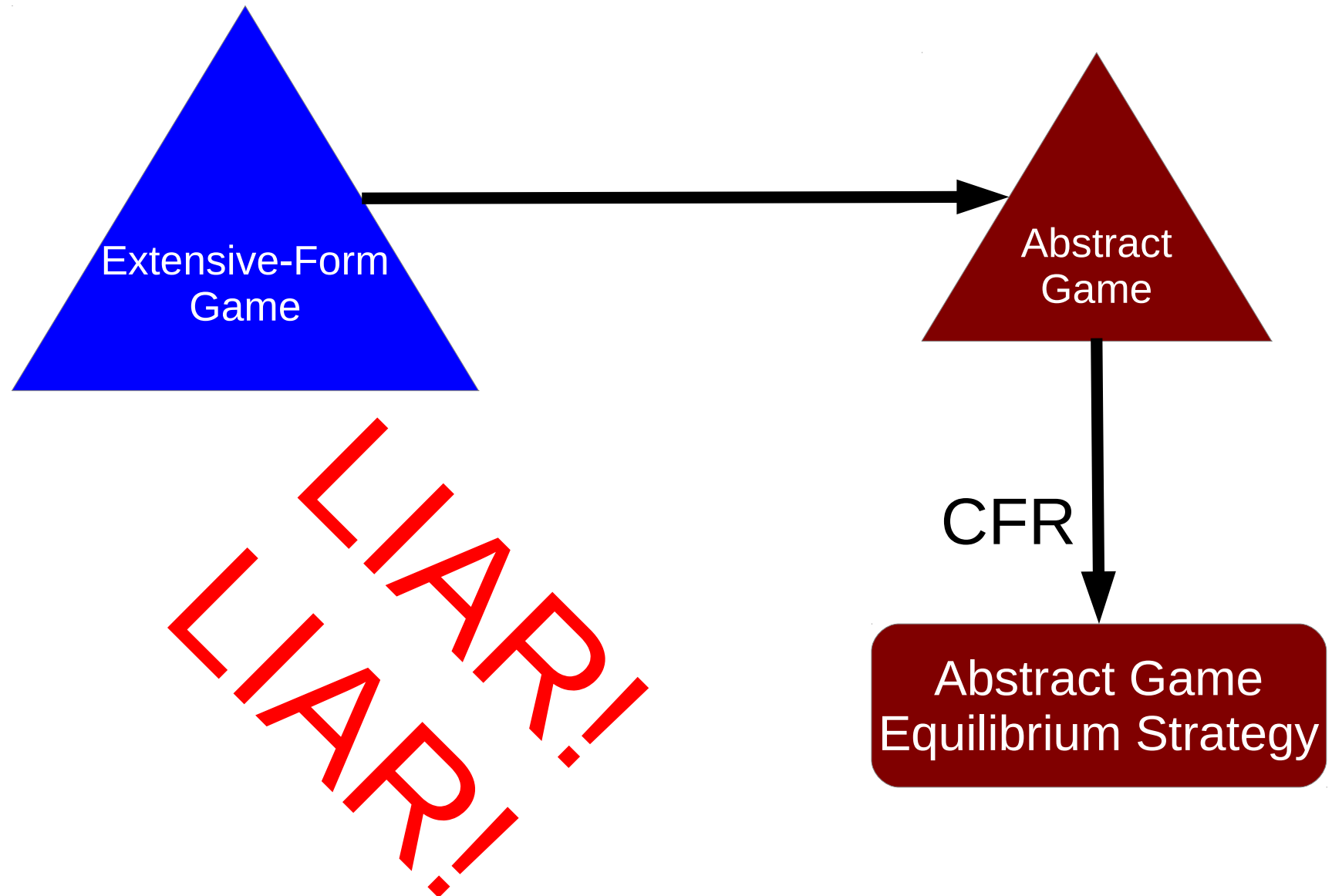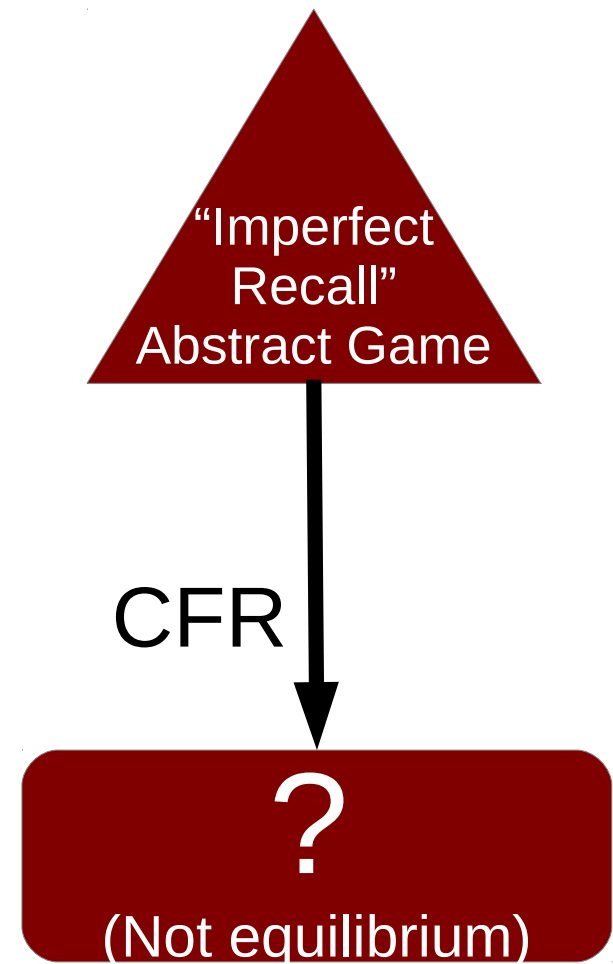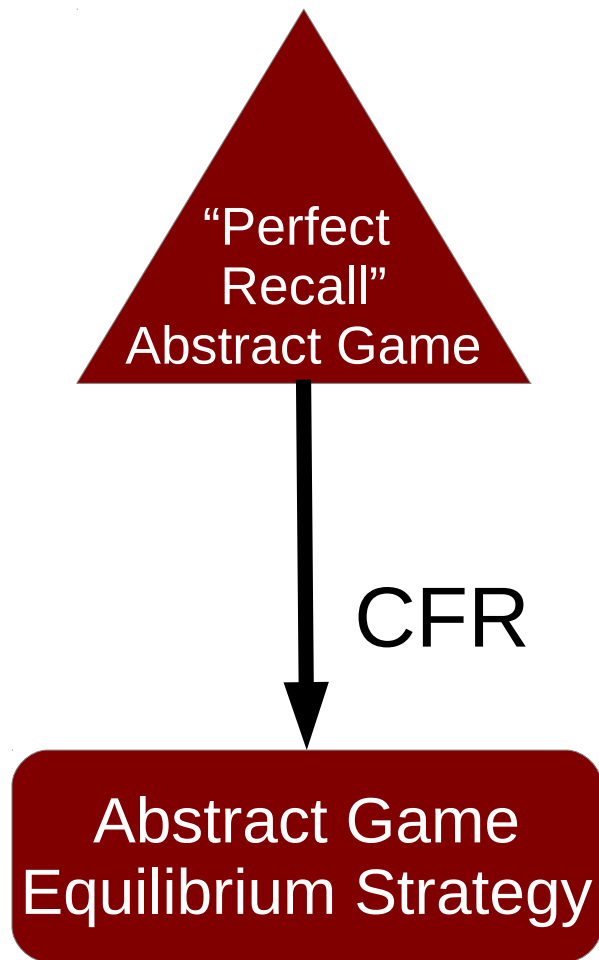
Imperfect Recall
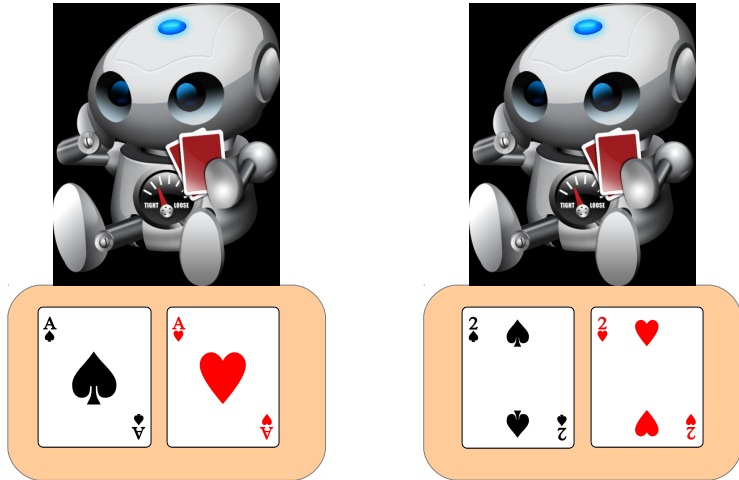
Imperfect Recall

Pre-flop

Imperfect Recall
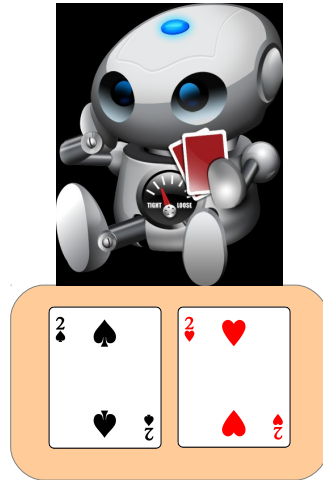
Pre-flop
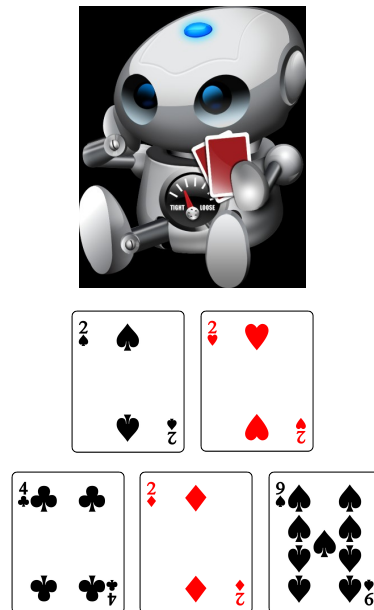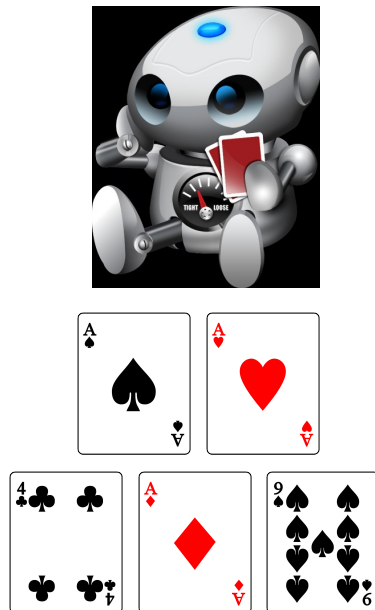
Flop

Imperfect Recall
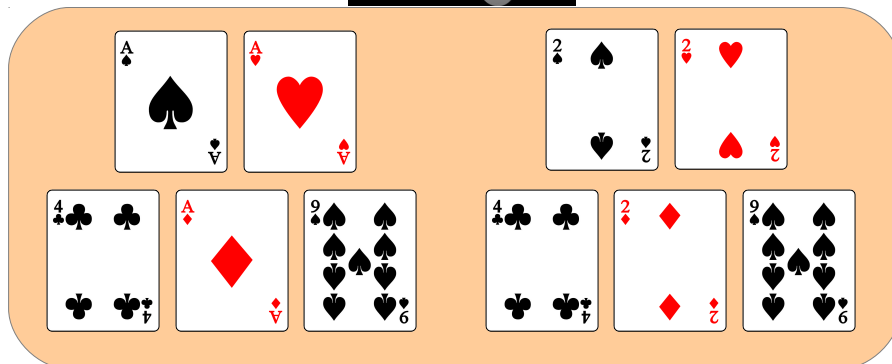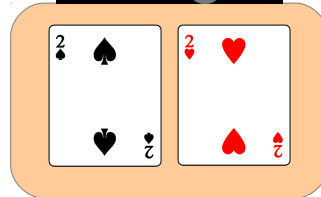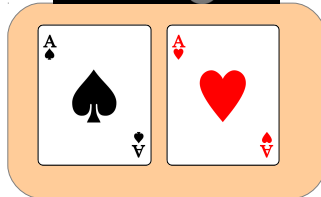
Imperfect Recall
Abstract Game

Imperfect Recall

Perfect Recall
Abstract Game

Imperfect Recall

Extensive-Form Game

Well-formed Imperfect Recall Game

CFR

New!

Abstract Game Equilibrium Strategy

- Unfortunately, does not prove convergence for our best poker abstractions.

- However, this was used in diabetes treatment research [Chen and Bowling, NIPS 2012].
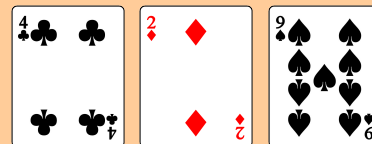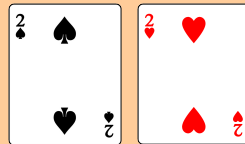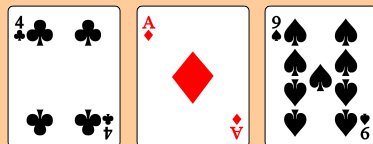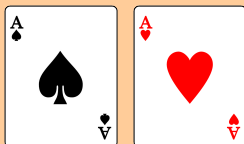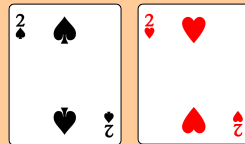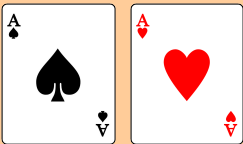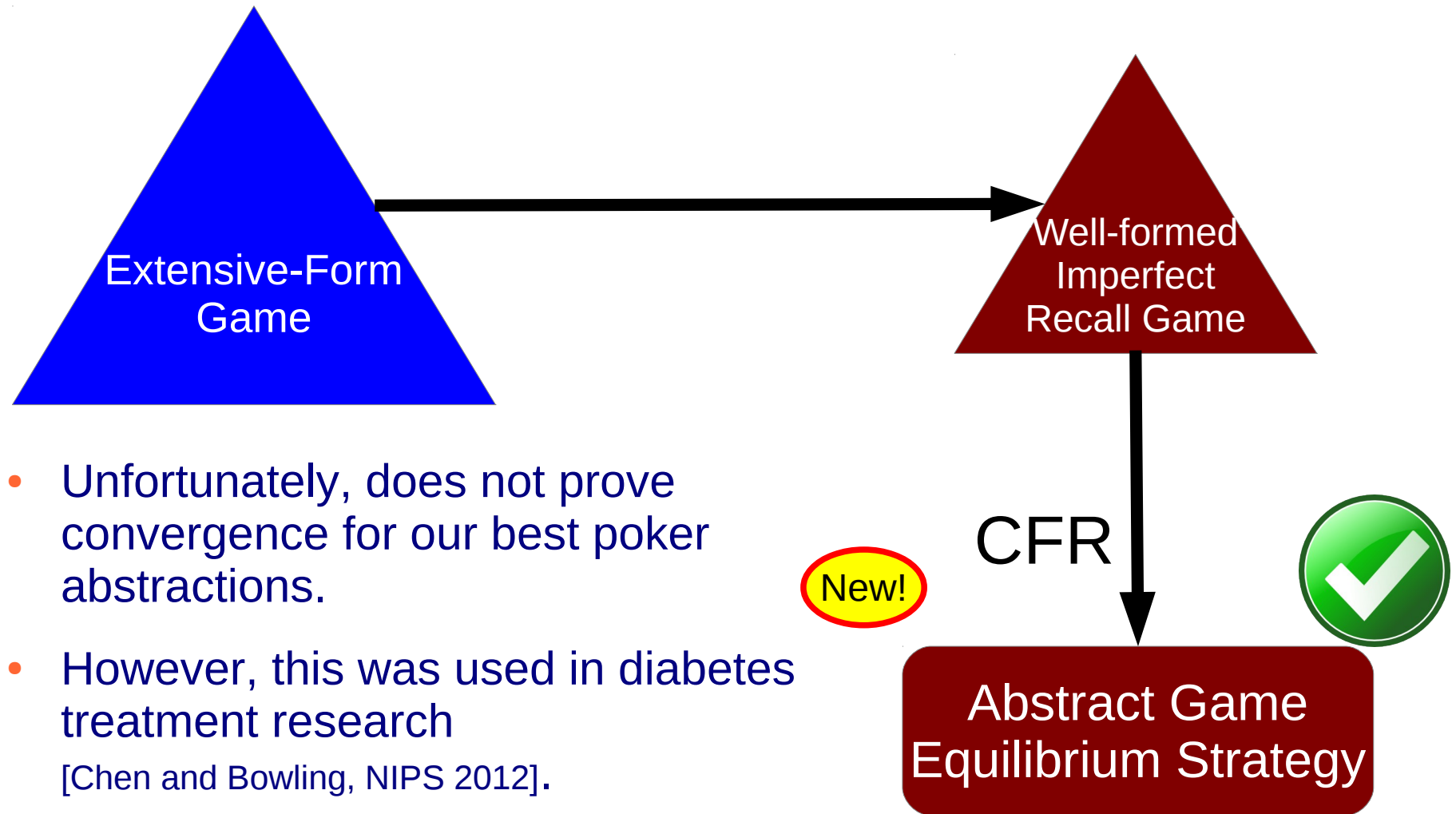
[Lanctot, G., Burch and Bowling, ICML 2012]

Outline of Presentation

- Background

  – Counterfactual Regret Minimization (CFR)

- Theoretical Advancements for CFR in:

  – Many player games

  – Imperfect recall games

- **CFR Speed-Ups**

- Tricks with Memory Limitations

- Conclusion + Future Work

CFR Speed-ups

CFR Speed-ups

- Each iteration, only update action probabilities at a sampled subset of states.
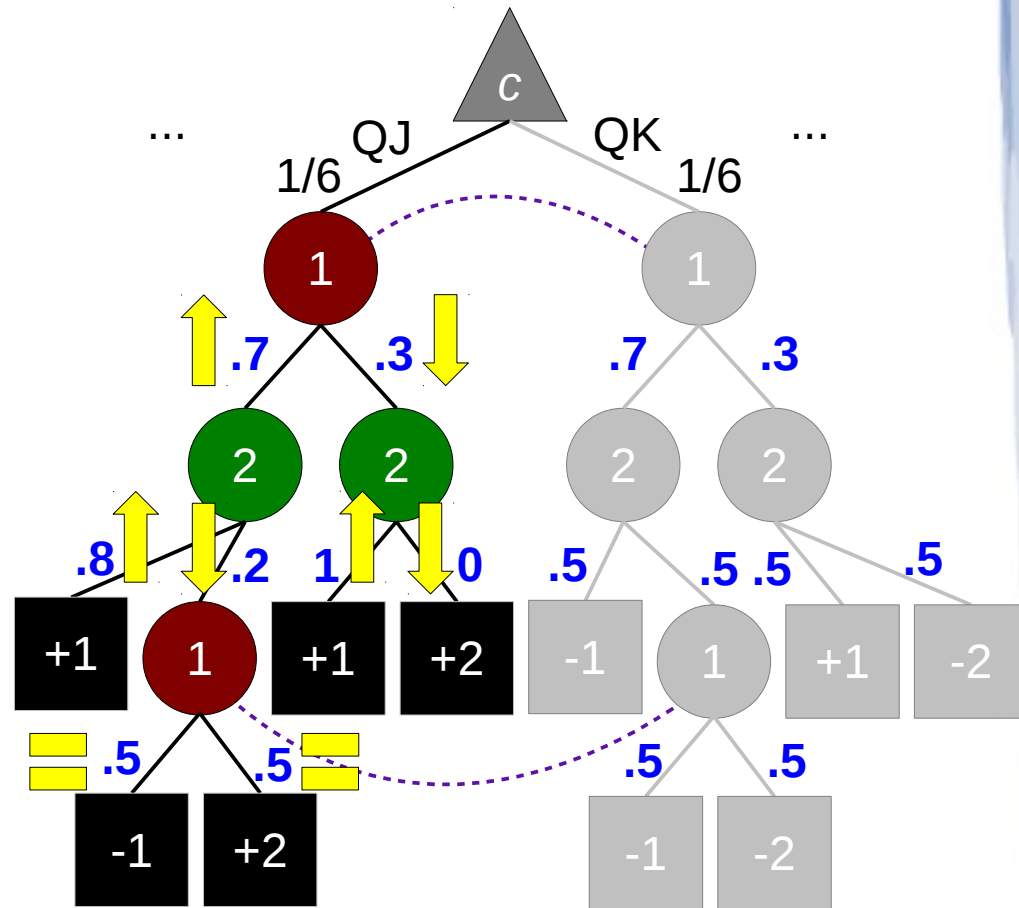
CFR Speed-Ups

Chance Sampling
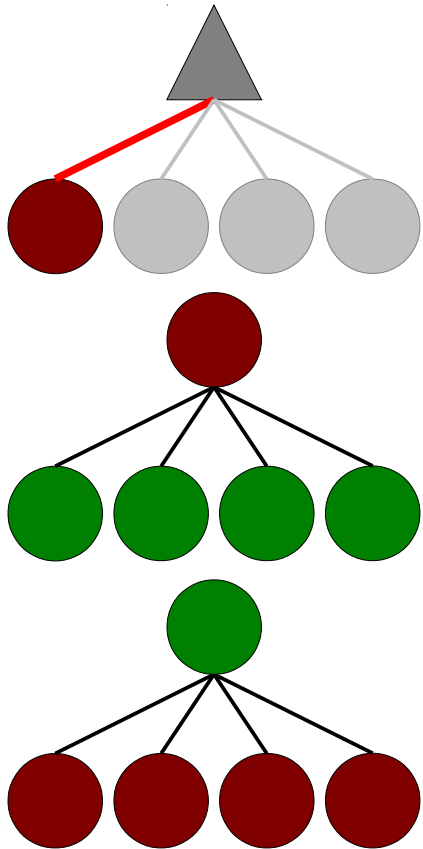
[Zinkevich *et al.*, NIPS 2007]

CFR Speed-Ups
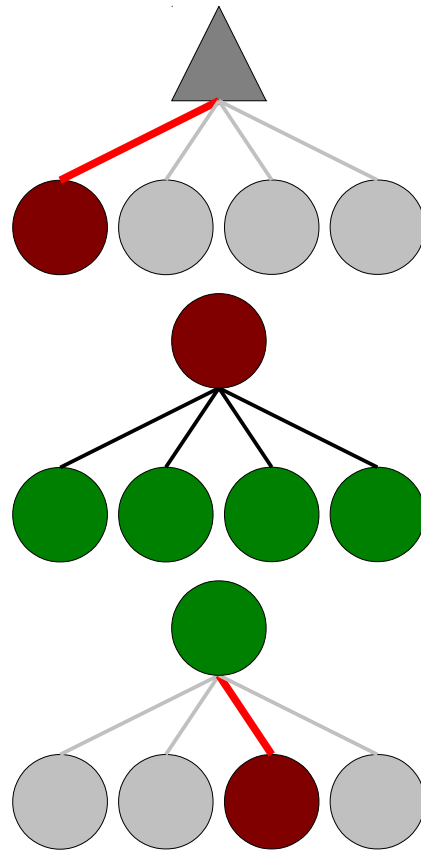
Chance Sampling | **External Sampling**

- Faster iterations
  - Use new strategies sooner
- Need more iterations
  - Good trade-off

[Zinkevich *et al.*, NIPS 2007]   [Lanctot *et al.*, NIPS 2009]
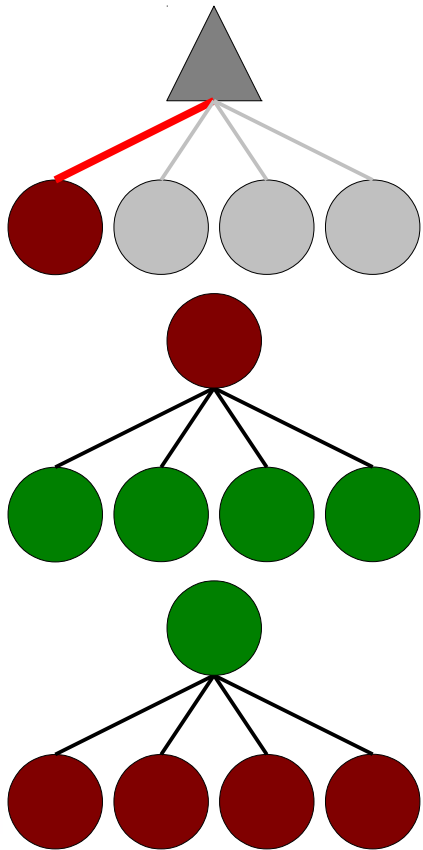
CFR Speed-Ups
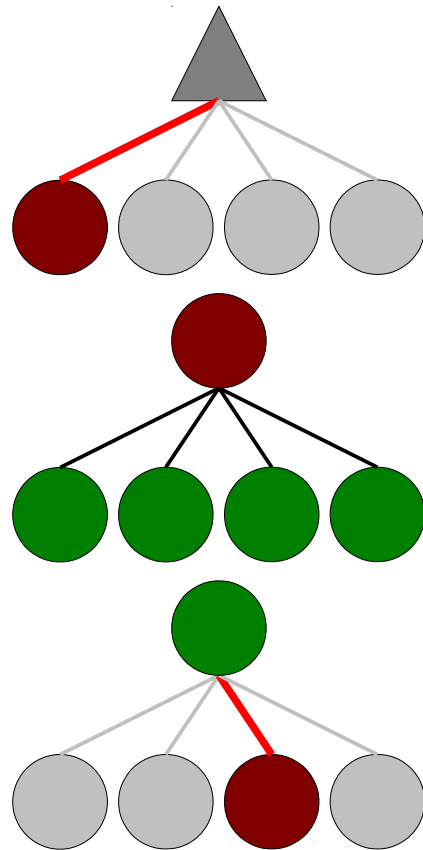
Chance Sampling | **External Sampling** | **Outcome Sampling**
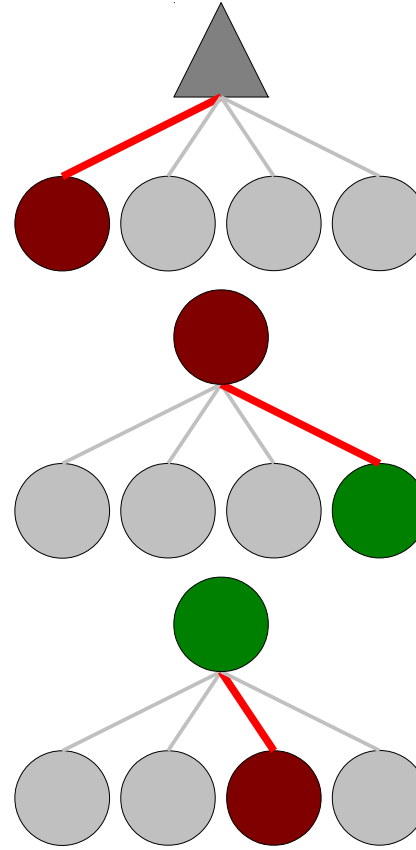
- Even faster iterations
- Even more iterations required
- Good trade-off?

[Zinkevich *et al.*, NIPS 2007]   [Lanctot *et al.*, NIPS 2009]

CFR Speed-Ups

2-round Heads-Up No-Limit Hold'em, 36 chips per player
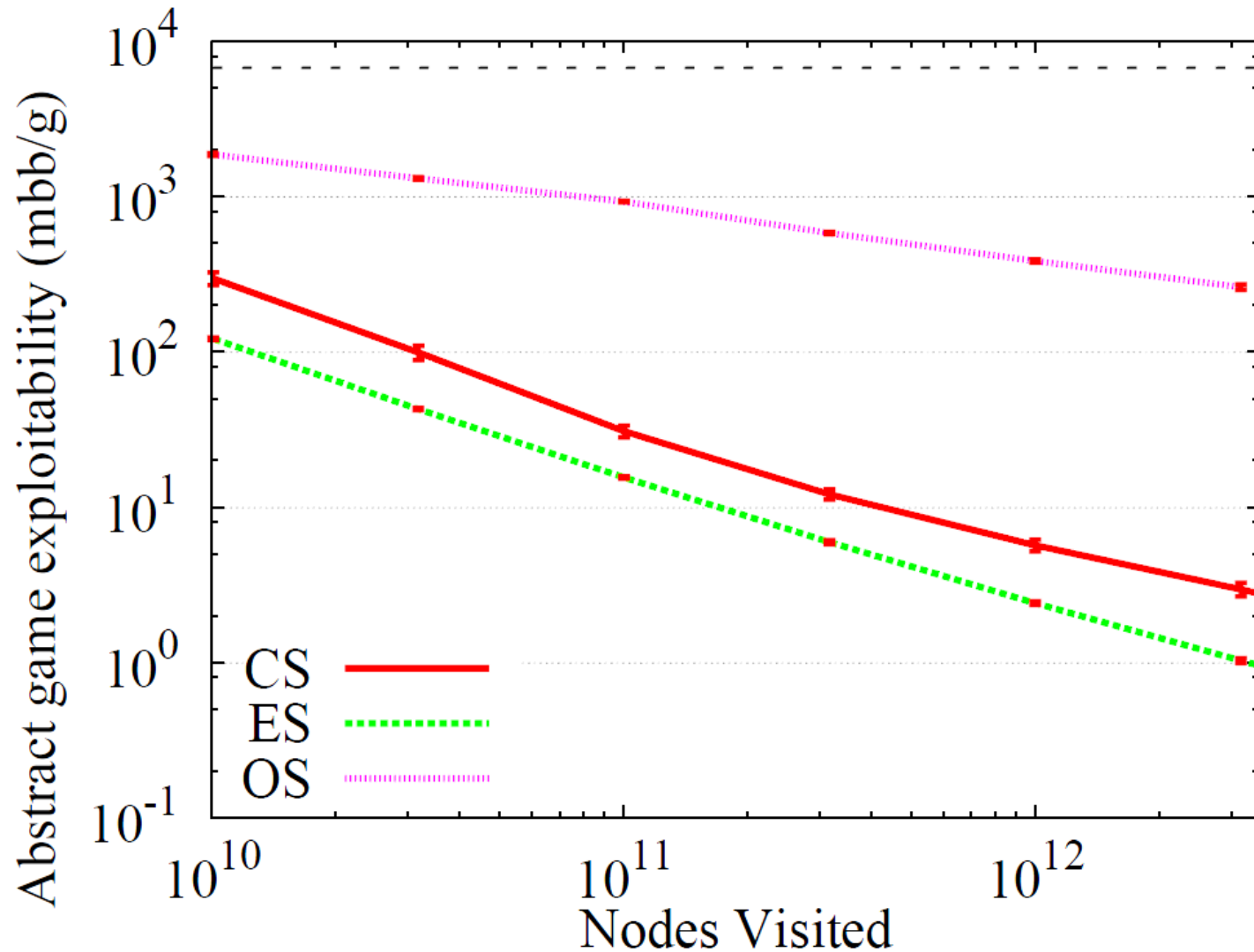
CFR Speed-Ups

- Outcome sampling introduces a lot of **variance**

- $T \leq C + K \cdot$ **Variance** [G. et al., AAAI 2012]

  New!

  – T = Iterations required to be

      "close enough" to equilibrium

  – C, K = Constants
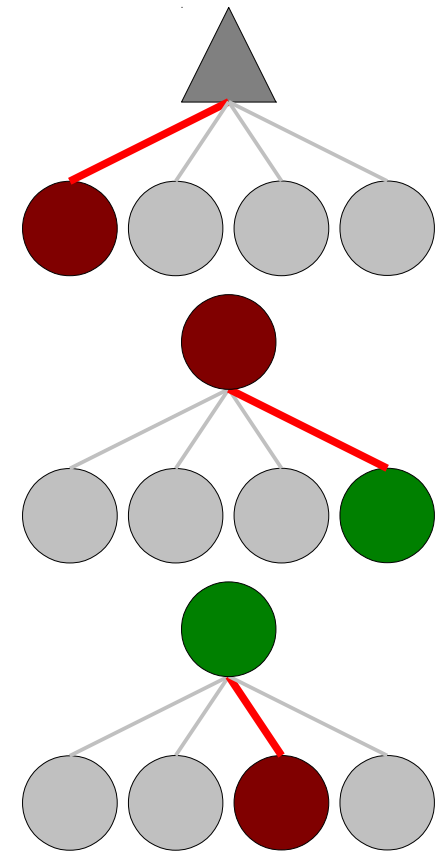
[G., Lanctot, Burch, Szafron and Bowling, AAAI 2012]

CFR Speed-Ups

Chance Sampling
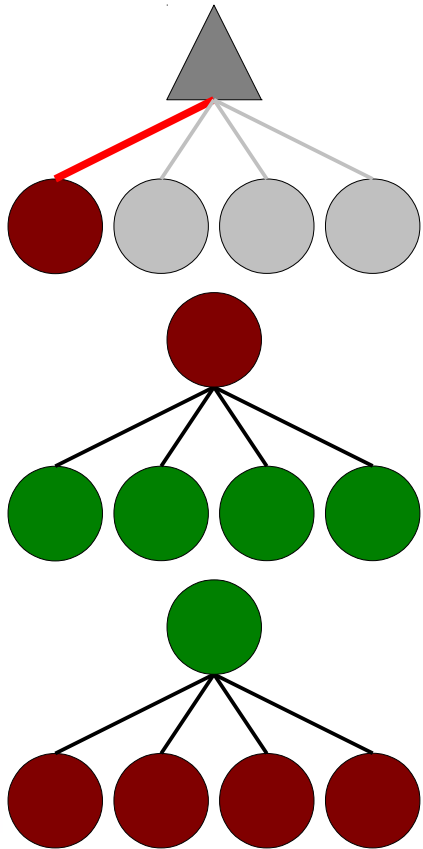
External Sampling

Outcome Sampling

Average Strategy Sampling

New!

Average Strategy

[Zinkevich *et al.*, NIPS 2007] [Lanctot *et al.*, NIPS 2009] [G. *et al.*, NIPS 2012]
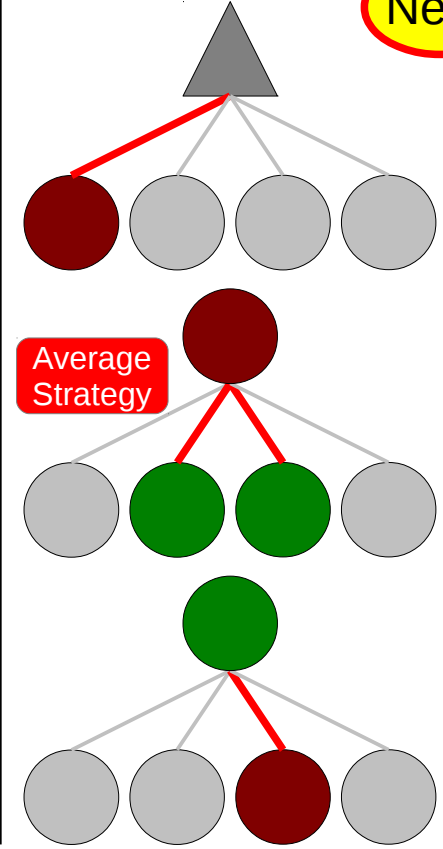
CFR Speed-Ups

2-round Heads-Up No-Limit Hold'em, 36 chips per player
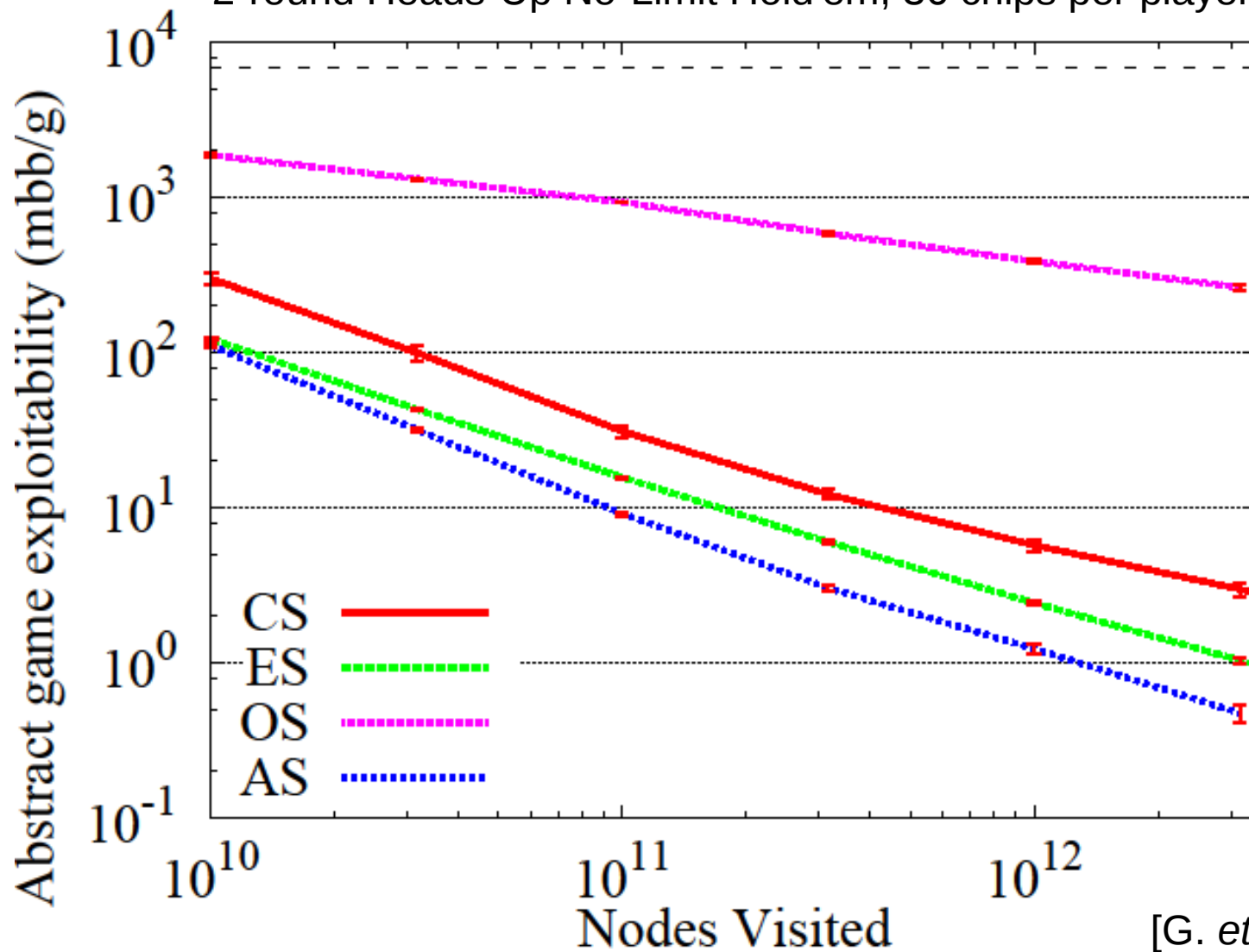
[G. *et al.*, NIPS 2012]

CFR Speed-Ups

2-round Heads-Up No-Limit Hold'em, *k* chips per player
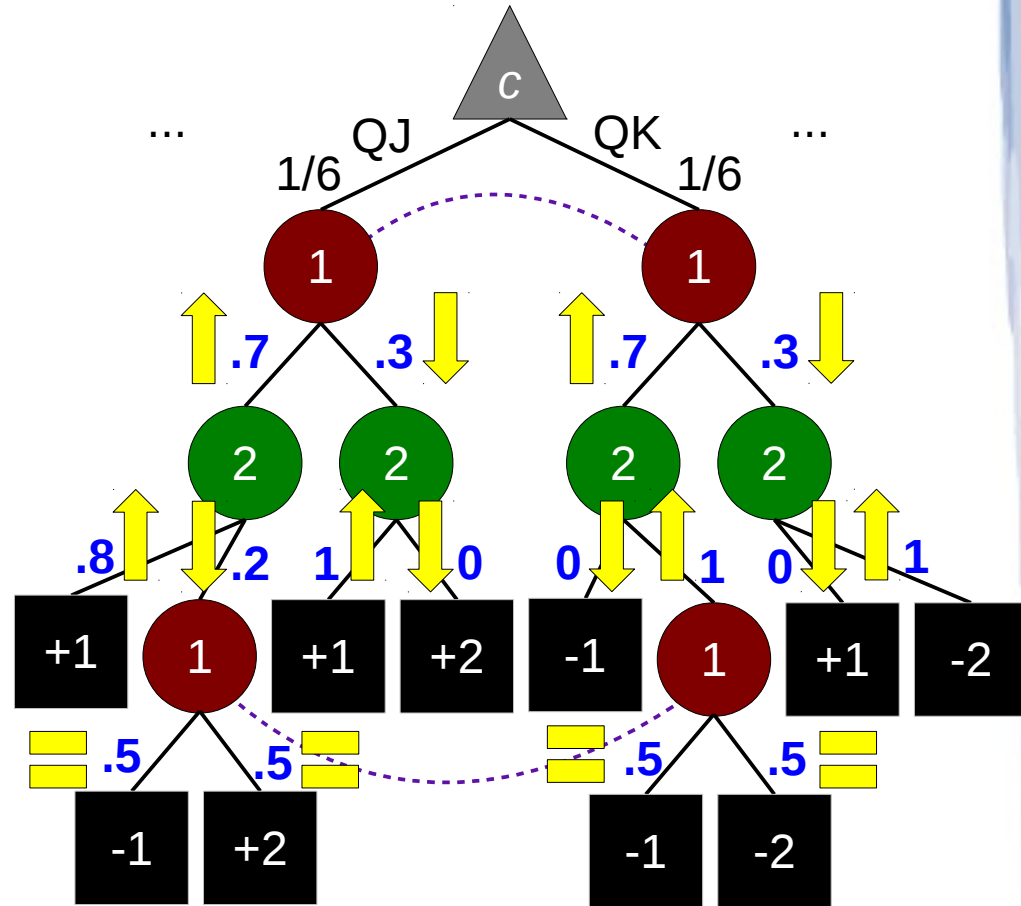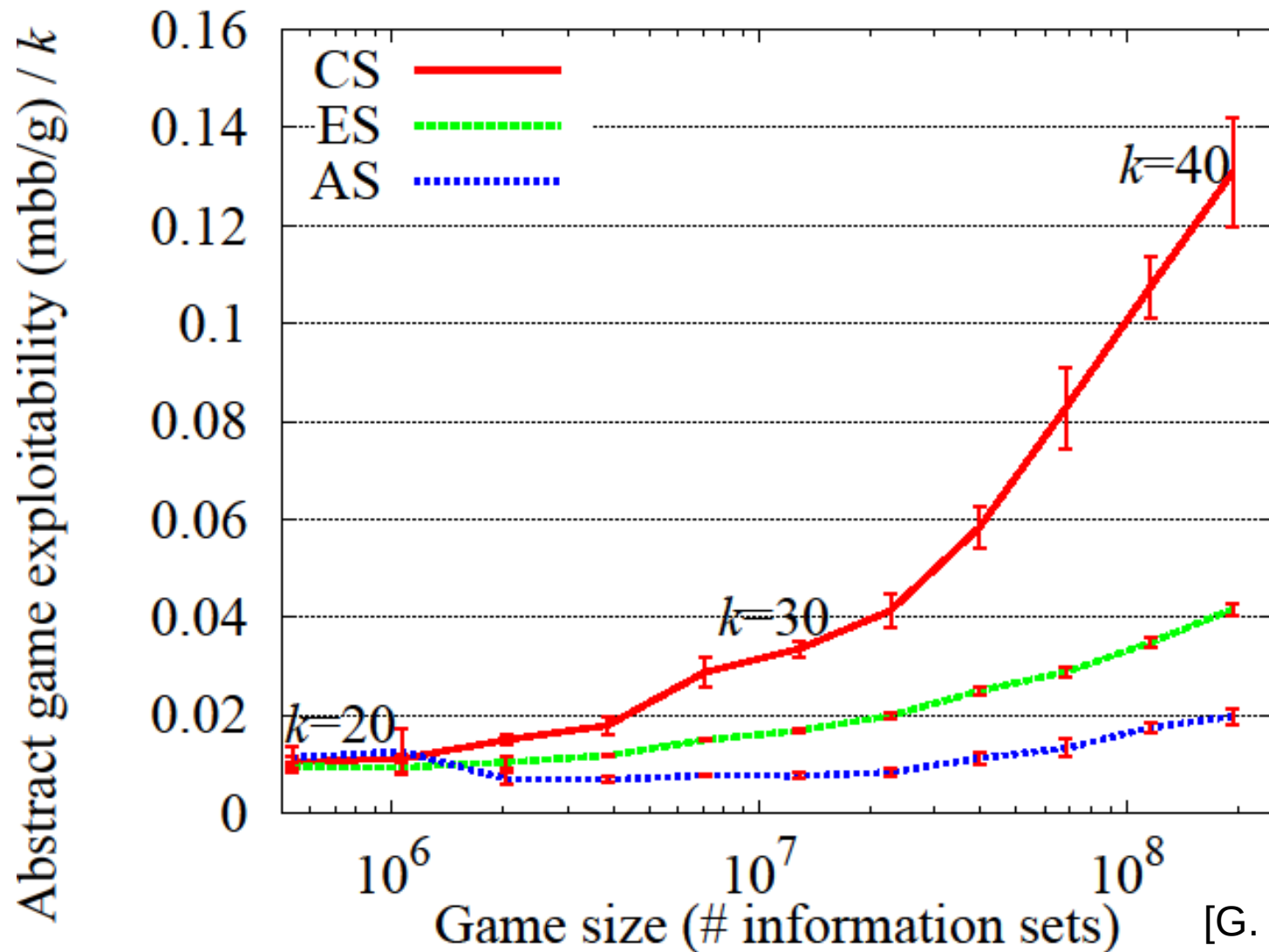

[G. *et al.*, NIPS 2012]

Outline of Presentation

- Background

    – Counterfactual Regret Minimization (CFR)

- Theoretical Advancements for CFR in:

    – Many player games

    – Imperfect recall games

- CFR Speed-Ups

- **Tricks with Memory Limitations**

- Conclusion + Future Work

Tricks with Memory Limitations

2-player Limit Texas Hold'em $\approx 10^{14}$

Abstract Game $\approx 10^{9}$

$\approx$ 59,000,000 "Turn" Deals

540,000 "Turn" Buckets

Tricks with Memory Limitations

2-player Limit Texas Hold'em $\approx 10^{14}$

Abstract Game $\approx 10^{9}$

$\approx 59,000,000$ "Turn" Deals

540,000 "Turn" Buckets

3-player Limit Texas Hold'em $\approx 10^{17}$

Abstract Game $\approx 10^{9}$

$\approx 59,000,000$ "Turn" Deals

540 "Turn" Buckets

Tricks with Memory Limitations

Tricks with Memory Limitations

3-Player Limit Texas Hold'em

≈ 59,000,000 "Turn" Deals

Abstract Game

540 "Turn" Buckets

2-player Experts

2-player Sub-games

3-player Limit Texas Hold'em Stitched Strategy

Abstract Game Strategy

540,000 "Turn" Buckets

- Generalizes 3 previous approaches

[Gibson and Szafron, NIPS 2011]

Tricks with Memory Limitations

Extensive-Form Game

Abstraction 1

Abstraction 2

Abstraction *K*

New!

"Frankenstein" Abstract Game

Full Game Strategy

Frankenstein-Game Strategy

[Gibson and Szafron, NIPS 2011]

Tricks with Memory Limitations

3-player Limit Texas Hold'em

18K "Turn" Buckets

1.53 Million "Turn" Buckets

New!

"Frankenstein" Abstract Game

Frankenstein-Game Strategy

3-player Texas Hold'em Strategy
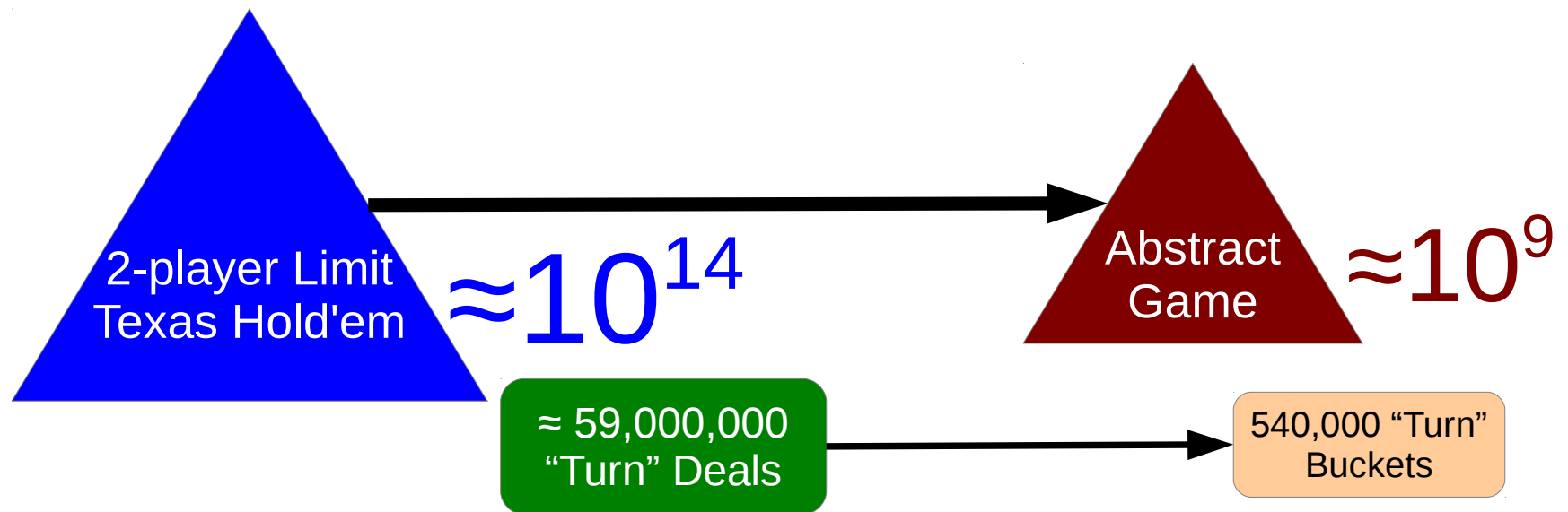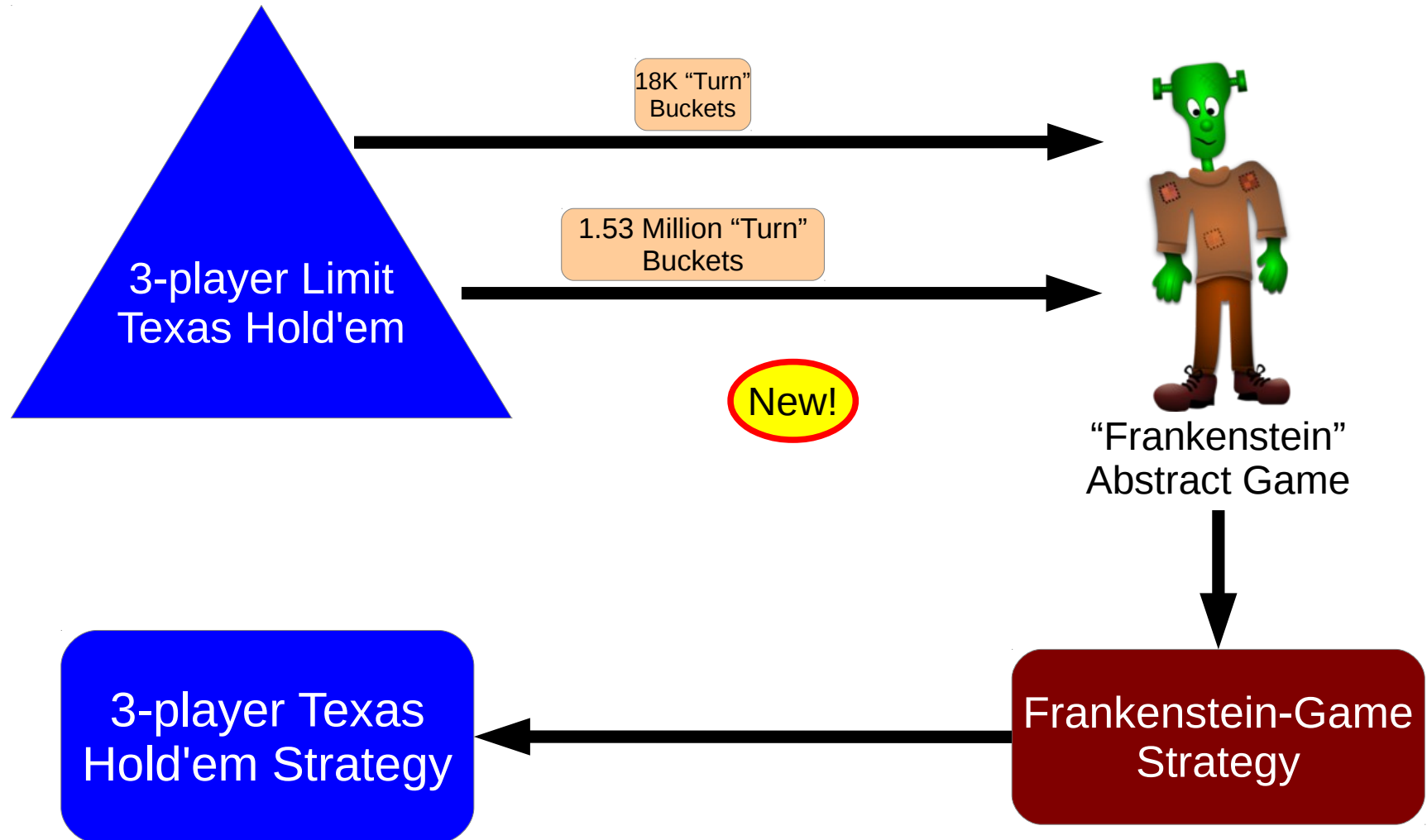
Tricks with Memory Limitations

Hyperborean3p Tournament
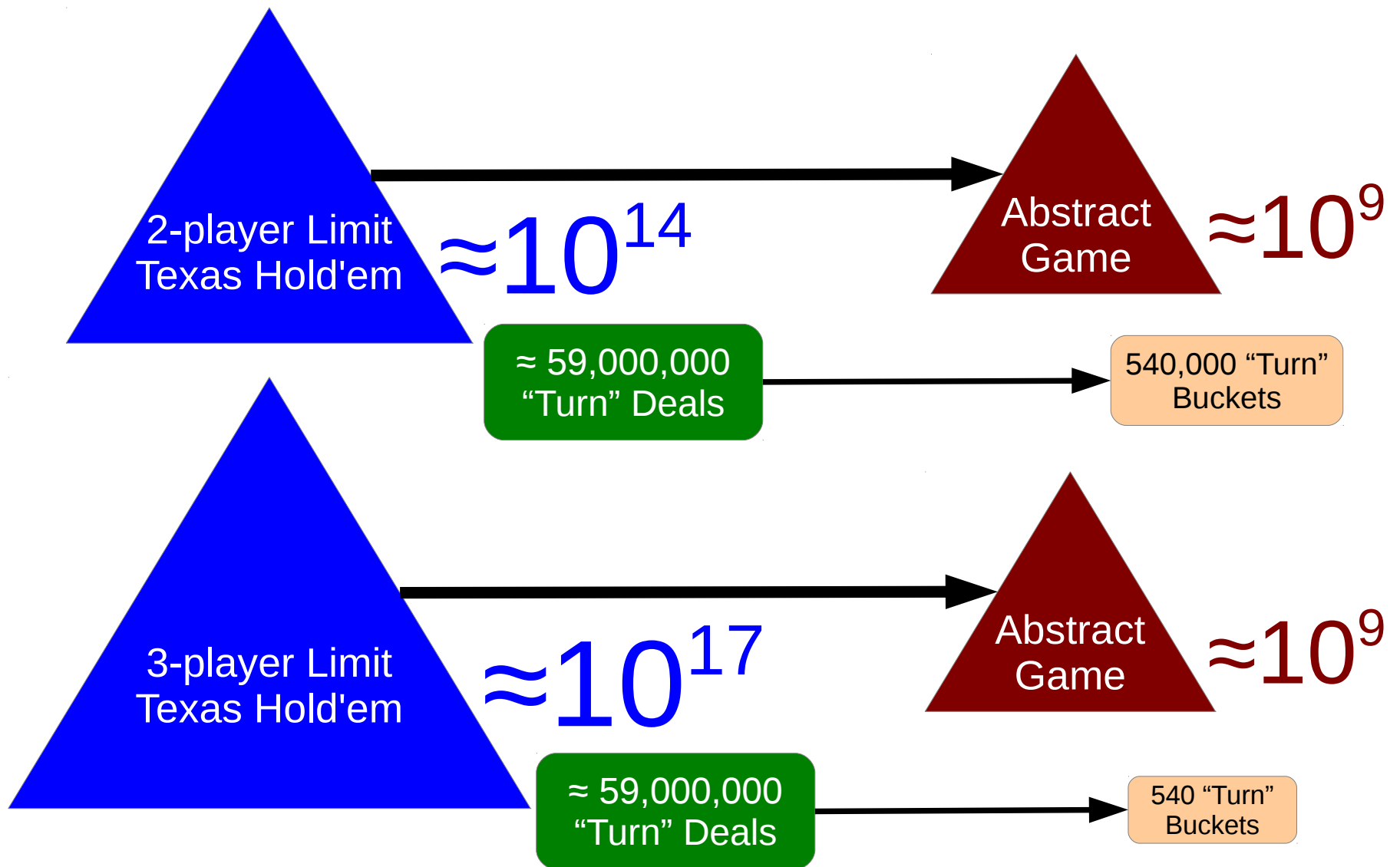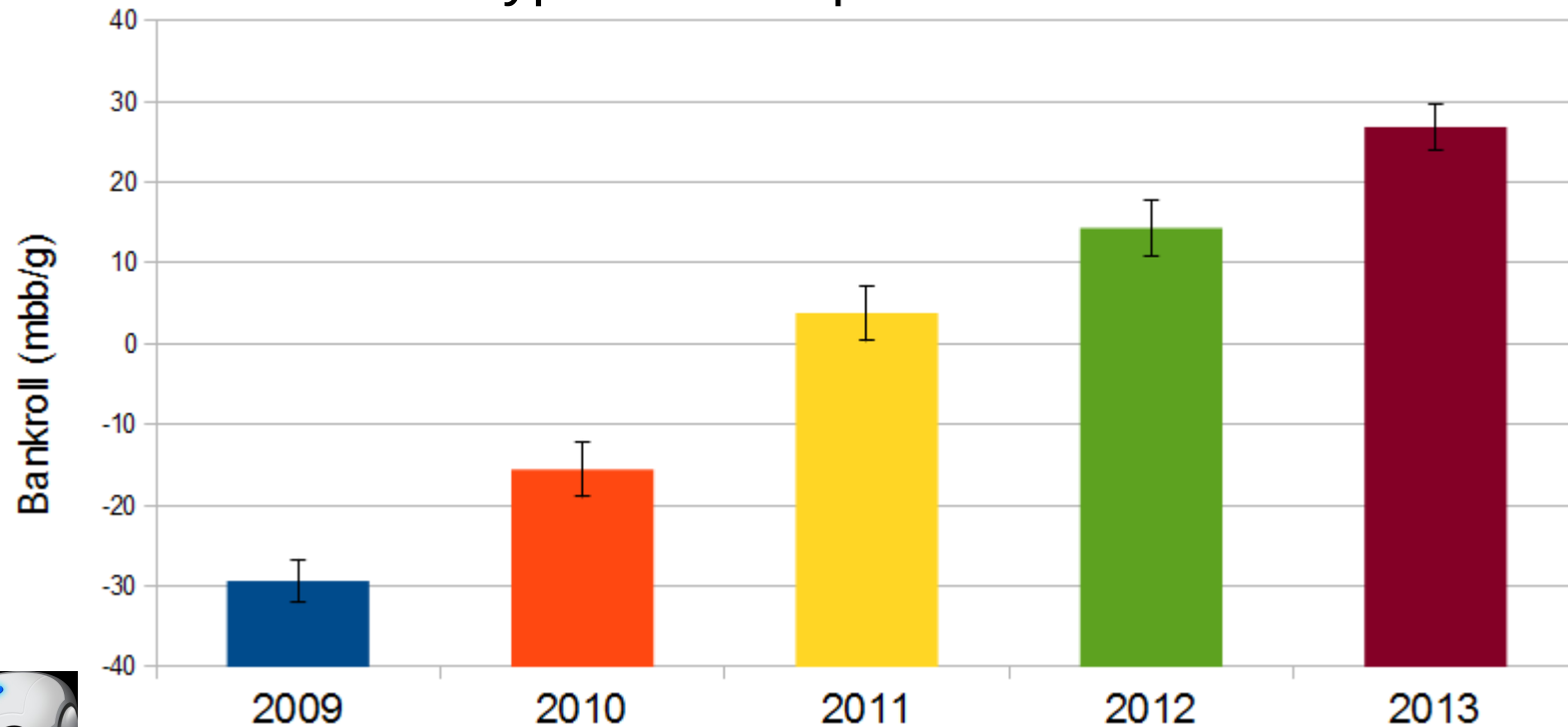
Outline of Presentation

- Background

  – Counterfactual Regret Minimization (CFR)

- Theoretical Advancements for CFR in:

  – Many player games

  – Imperfect recall games

- CFR Speed-Ups

- Tricks with Memory Limitations

- **Conclusion + Future Work**

Conclusion

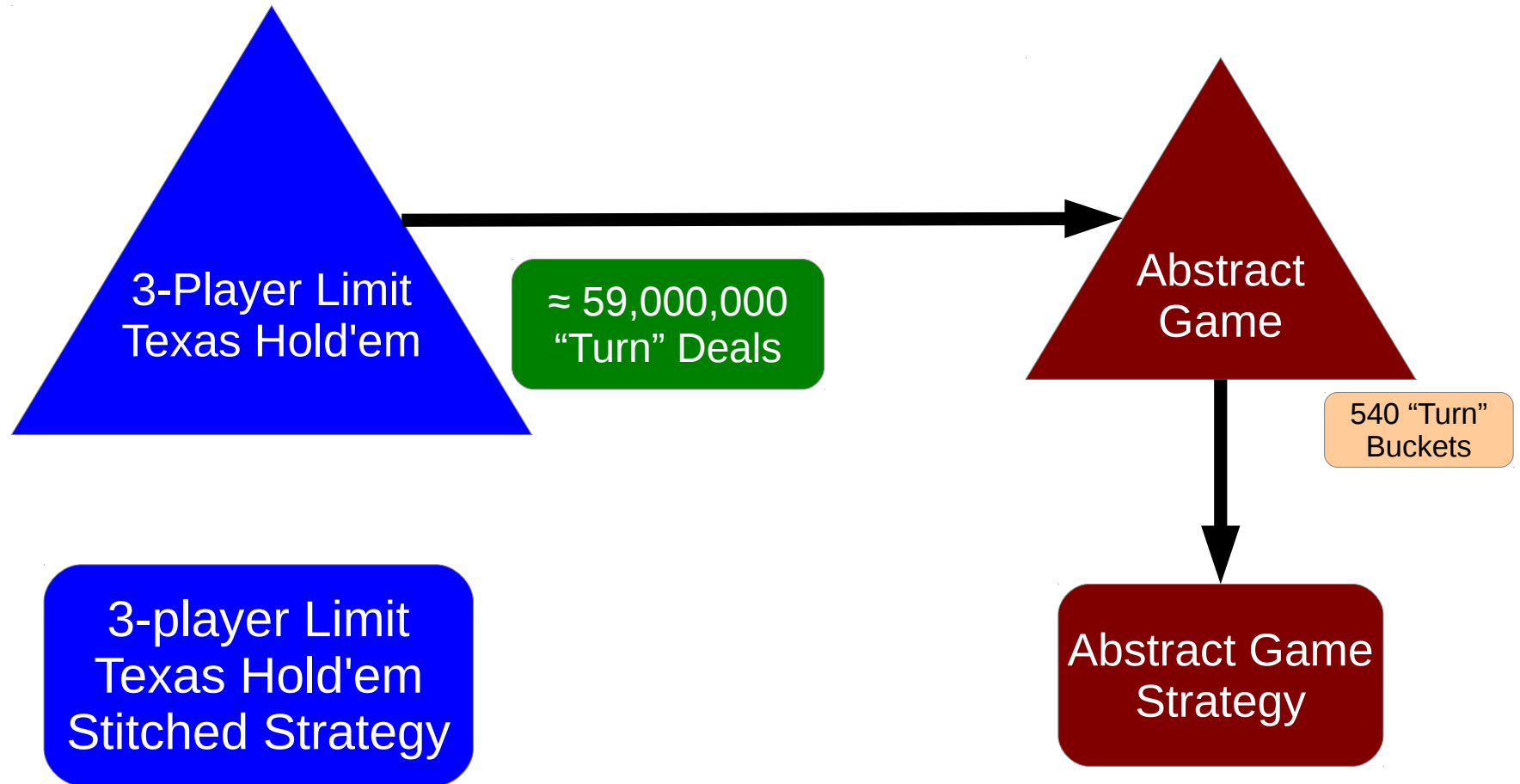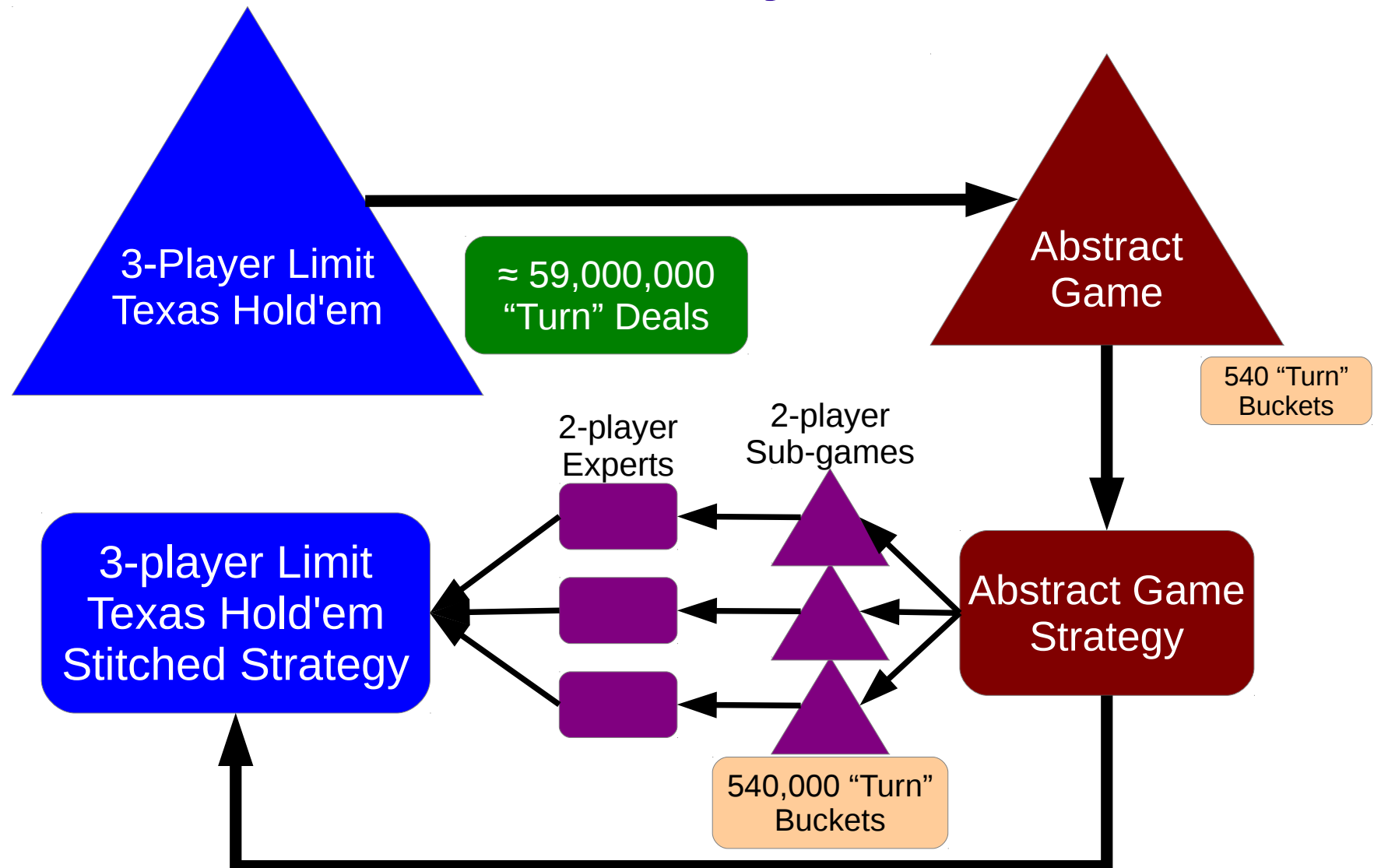- We have made the following contributions:
  - First set of theoretical properties for CFR in:
    - games with more than 2 players
    - imperfect recall games
  - Theoretical and practical improvements for making CFR go faster
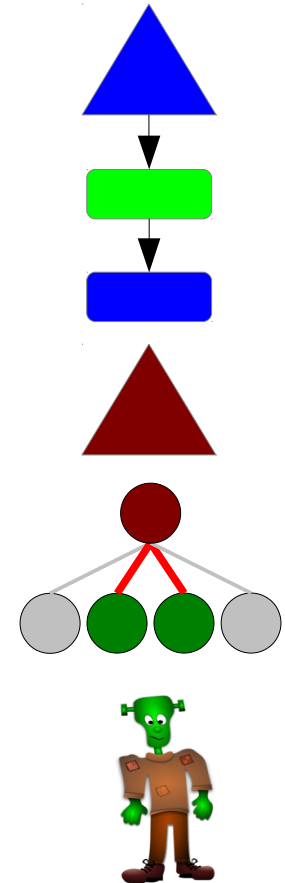 - Techniques for dealing with limited memory
- This research has led to the development of the strongest 3-player limit Texas hold'em strategies in the world.

Future Work

- Opponent modelling
  - "On-line CFR"
- 10-player Texas Hold'em
  - Ultimately challenge humans for the World Series of Poker

Thanks for Listening!

- Email: richard.g.gibson@gmail.com

- Website: http://cs.ualberta.ca/~rggibson/

- Twitter: @RichardGGibson

Clip art images used in this presentation can be found at clker.com