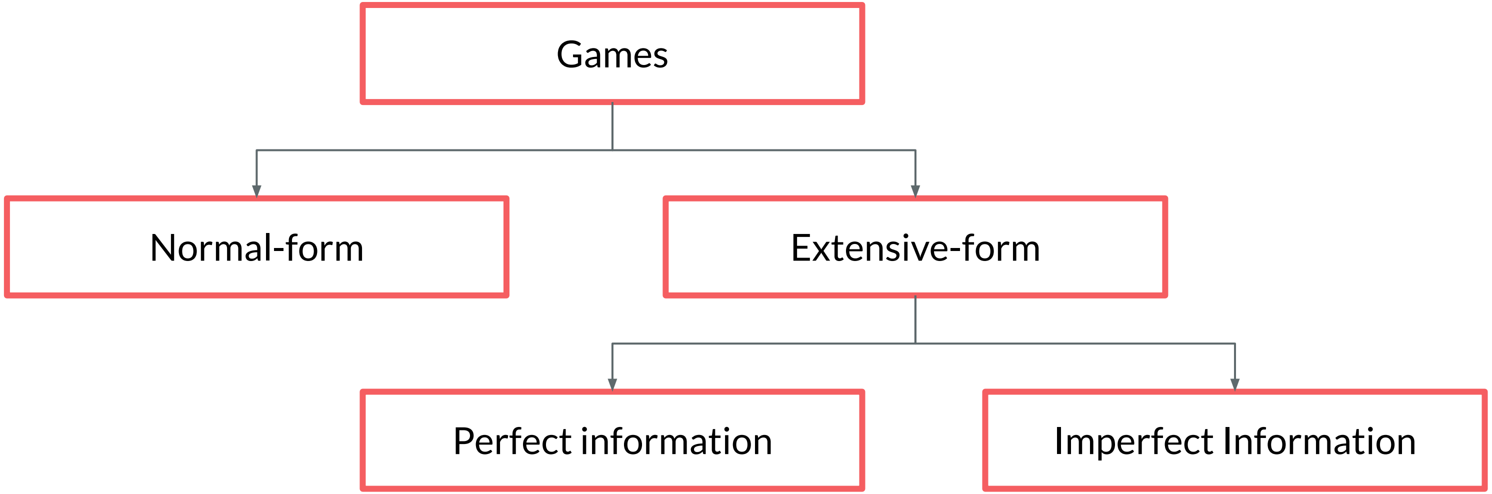
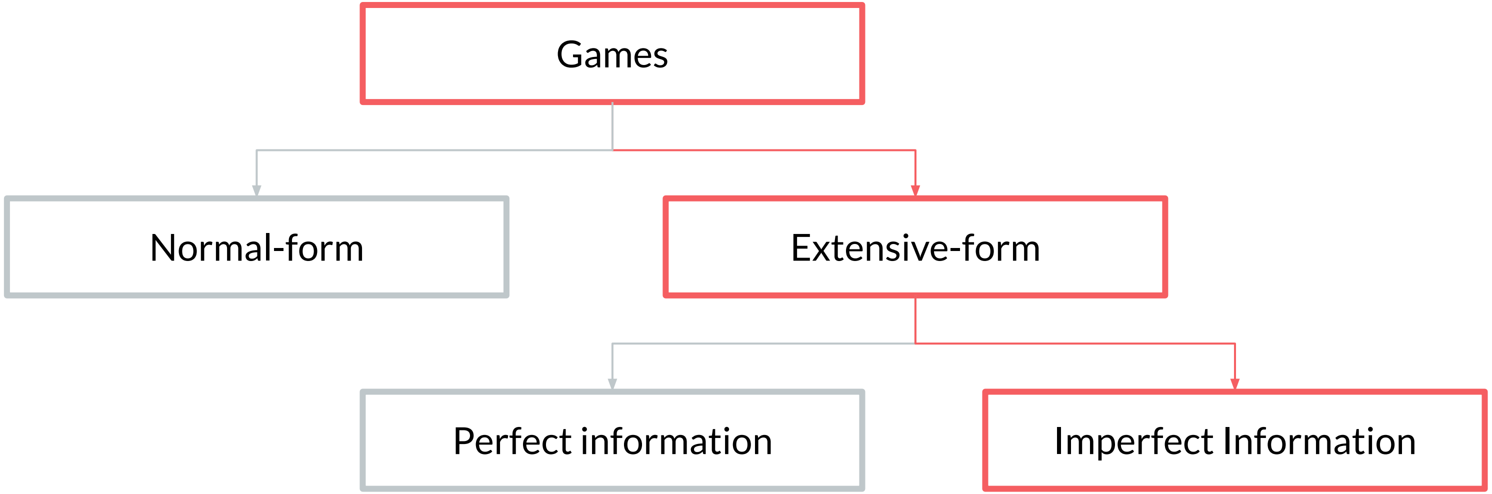
Learning Algorithms for Regret Minimization

Elaheh Toulabinejad, Mahshid Rahmani Hanzaki

Introduction


Games
Normal-form
Extensive-form
Perfect information
Imperfect Information


Games
Normal-form
Extensive-form
Perfect information
Imperfect Information

Linear Programming

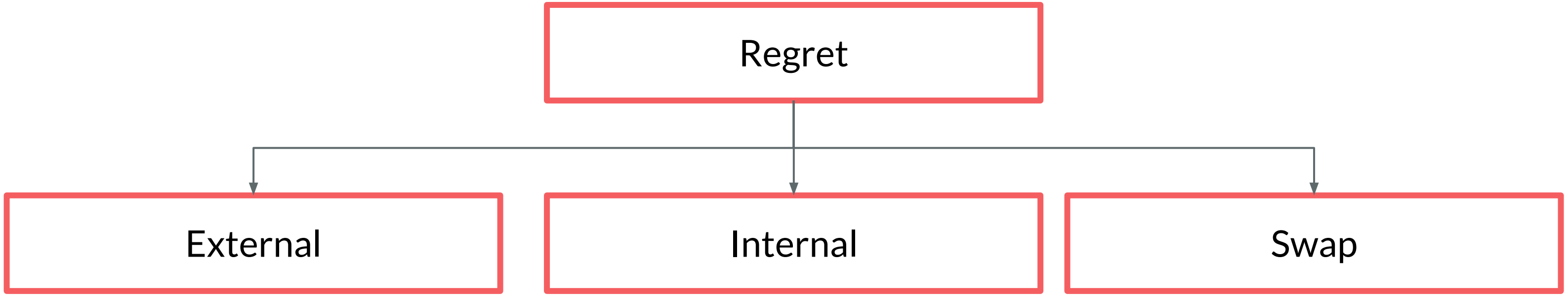
Normal-form Linear Programming

- One of earliest methods
- Convert to normal-form game
- Use Linear Programming (LP)
- Number of strategies exponential to the number of information sets

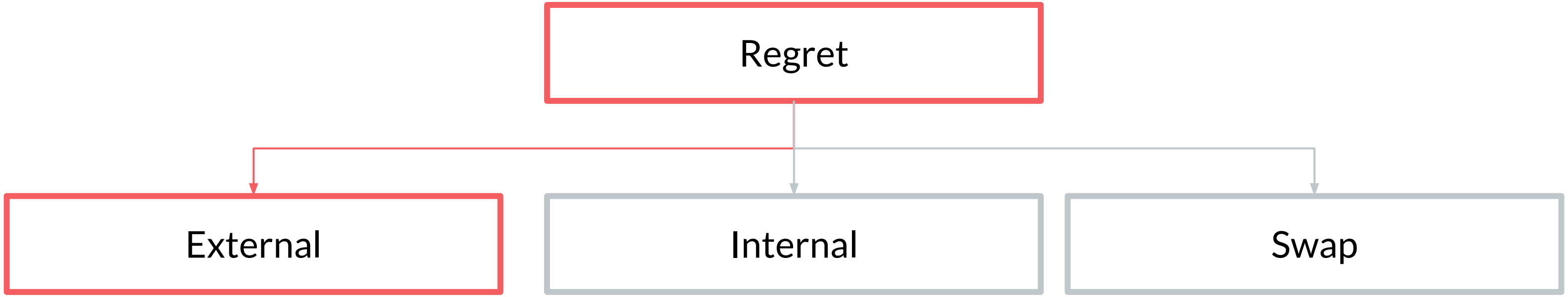
Sequence-form Linear Programming

- Modern era of solving Imperfect information game
- Introduced sequence-form representation of strategy
- Used LP for extensive-form
- Polynomial as size of game representation

Regret Minimization



Regret Minimization



External Regret Minimization in Extensive-form Games With Imperfect Information

Counterfactual Regret Minimization 2007

Counterfactual Regret

- Break down total regret into smaller parts for each information set.
- Reduce counterfactual regret separately for each set.
- Total regret is bounded by the sum of all counterfactual regrets.
- Lowering immediate counterfactual regret reduces overall regret.
- Finding an approximate Nash equilibrium by minimizing immediate counterfactual regret.

Algorithm

- Uses minimizing regret in self-play to calculate Nash equilibrium
- On each iteration:
  - traverses the entire game tree
  - updates the regrets for every infostate in the game according to policy profile
  - Defines a new policy based on these regrets
- Average of these policies converges to an approximate Nash equilibrium in two player zero-sum games

Abstraction

- Definition: merging information sets
    - In terminology of poker: grouping card sequences
- Goal: reducing the number of information sets for each player to a tractable size

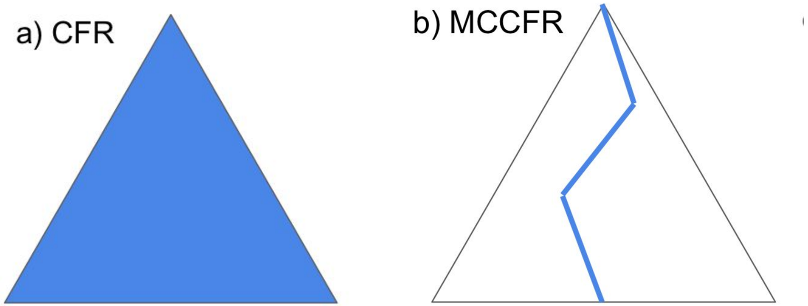
CFR Problem

traverse the entire tree on each iteration

Monte Carlo Sampling for Regret Minimization in Extensive Games 2009

Monte Carlo CFR (MCCFR)

- A family of domain independent algorithms
- Same regret updates as CFR on expectation
- Approximate equilibrium using self-play
- Increase in number of iterations
    - Constant-factor increase
- Reduced cost on each iteration
    - Order reduction
- Introduce two methods of Sampling
    - External Sampling
    - Outcome sampling



External Samping vs. Outcome Sampling

External Sampling

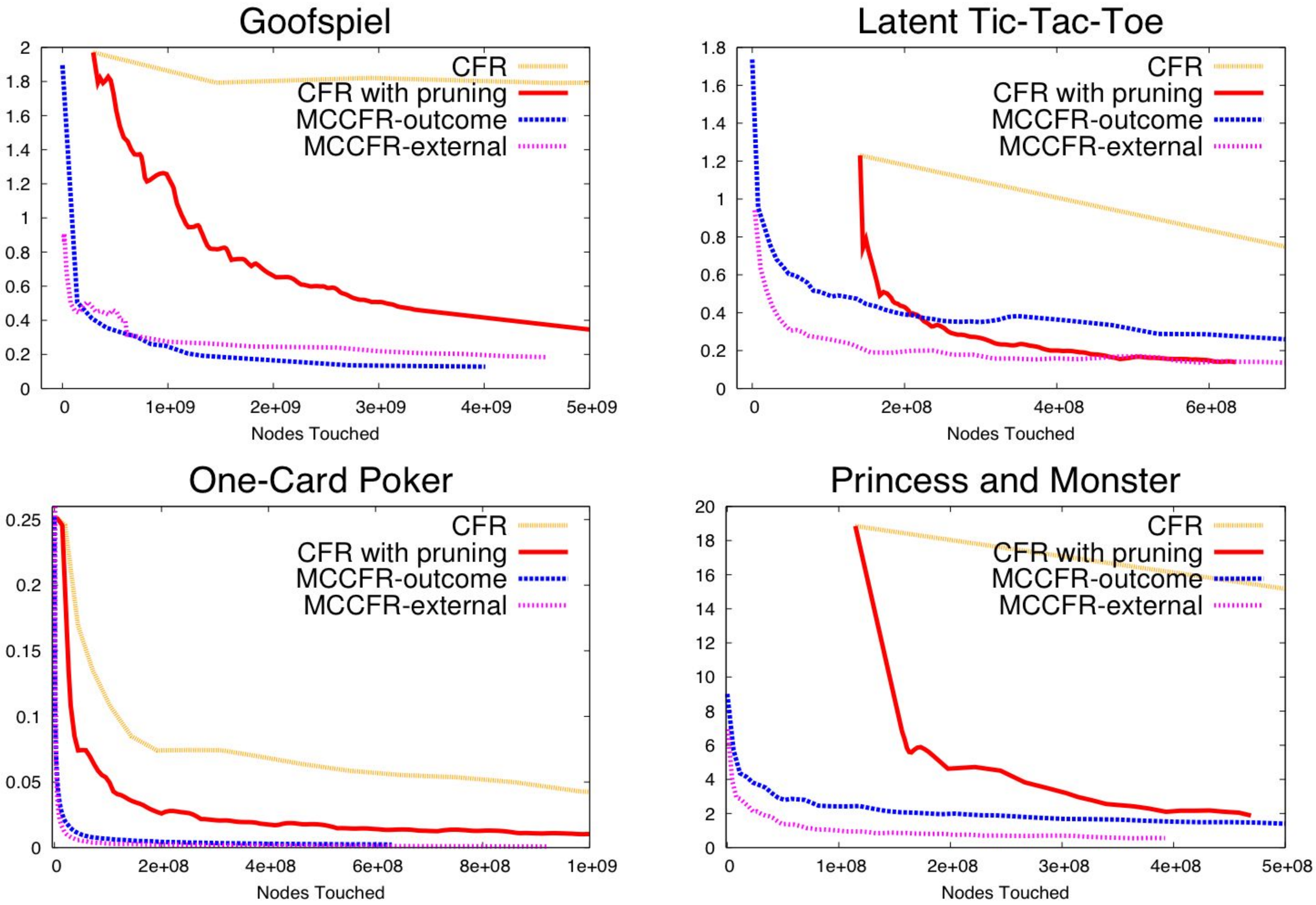
- Sample chance nodes (choices external to player and player actions

Outcome Sampling

- Single playing of game sampled
- Each sample only contains one outcome

Pruning

- Prune entire subtree if the other player has no probability of reaching



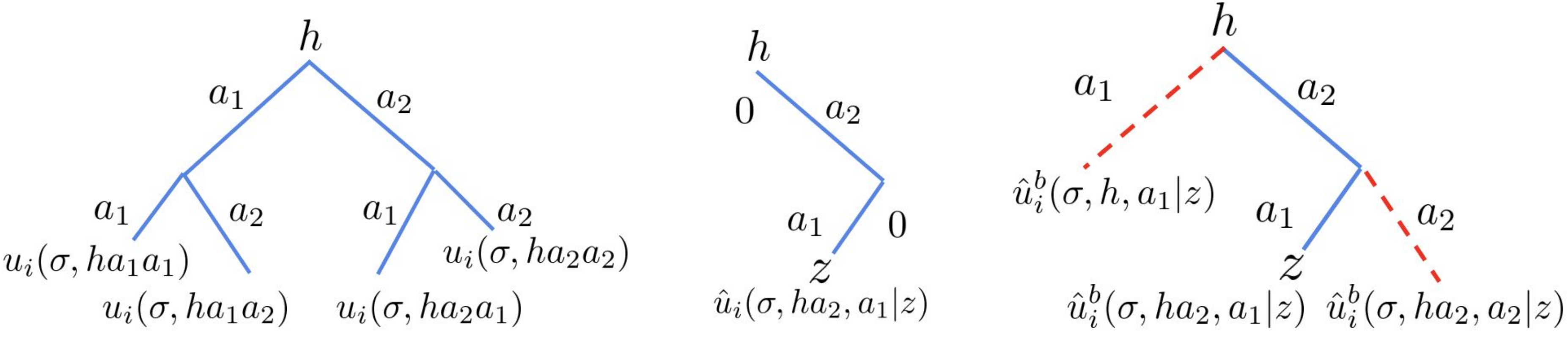
Goofspiel
CFR
CFR with pruning
MCCFR-outcome
MCCFR-external
Nodes Touched
Latent Tic-Tac-Toe
CFR
CFR with pruning
MCCFR-outcome
MCCFR-external
Nodes Touched
One-Card Poker
CFR
CFR with pruning
MCCFR-outcome
MCCFR-external
Nodes Touched
Princess and Monster
CFR
CFR with pruning
MCCFR-outcome
MCCFR-external
Nodes Touched

Main Problem with Sampling Methods

High variance

Variance Reduction in Monte Carlo Counterfactual Regret Minimization (VR-MCCFR) for Extensive Form Games Using Baselines 2019

VR-MCCFR

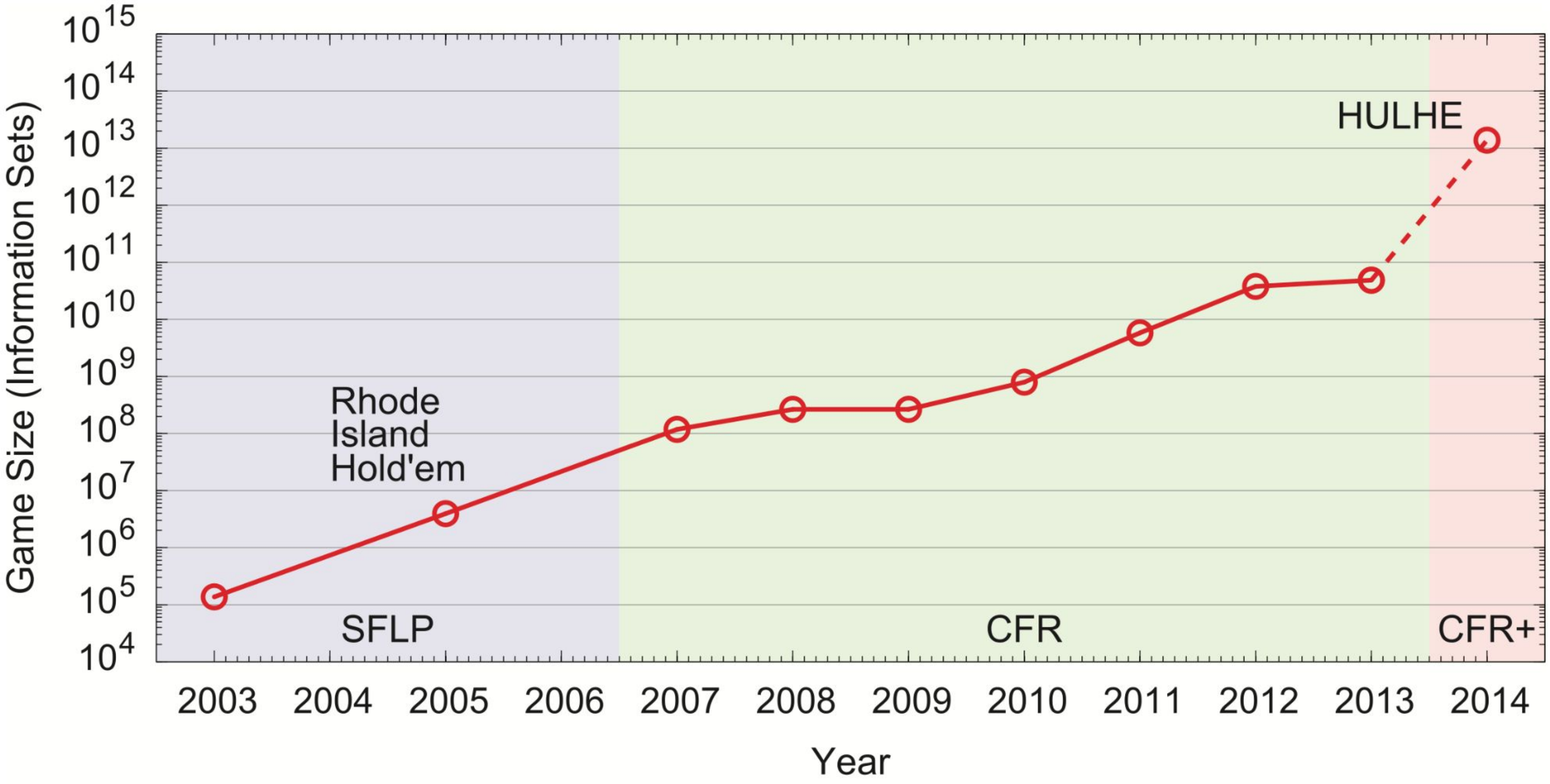


(a) CFR (b) MCCFR (c) VR-MCCFR

Unconventional Approach

CFR+

- Challenges of CFR in larger games
  - Memory challenge
    - Should store the accumulated regret values for each information set
- Use CFR+ for the computation problem
- CFR generally uses sampling to update on each iteration
- CFR+
  - Exhaustive search
  - Regret-matching+
  - Actions will be chosen immediately after proving useful
- CFR+ is the state-of-the-art for solving large imperfect information games



Game Size (Information Sets)
10^{15}
10^{14}
10^{13}
10^{12}
10^{11}
10^{10}
10^{9}
10^{8}
10^{7}
10^{6}
10^{5}
10^{4}
HULHE
Rhode Island Hold'em
SFLP
CFR
CFR+
2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014
Year

Double Neural Counterfactual Regret Minimization 2018

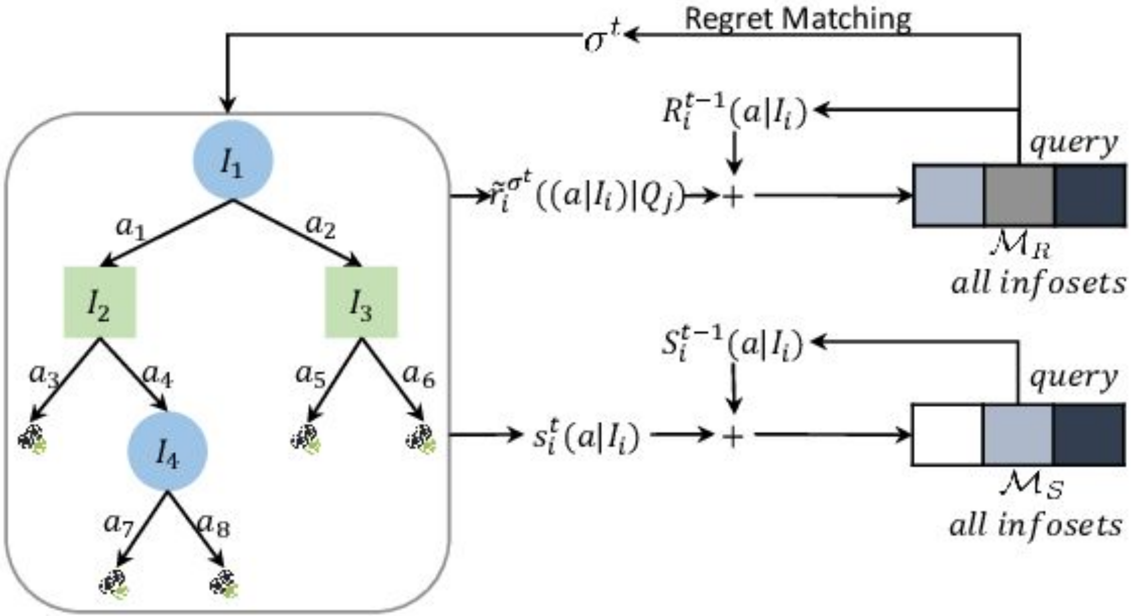
CFR Problems

- Only works for discrete state and action spaces using a tabular representation
  - Cannot directly be applied to large games
  - Cannot improve by starting from poor strategy profile
- Need to traverse the entire game tree
  - Cannot handle large games with limited memory

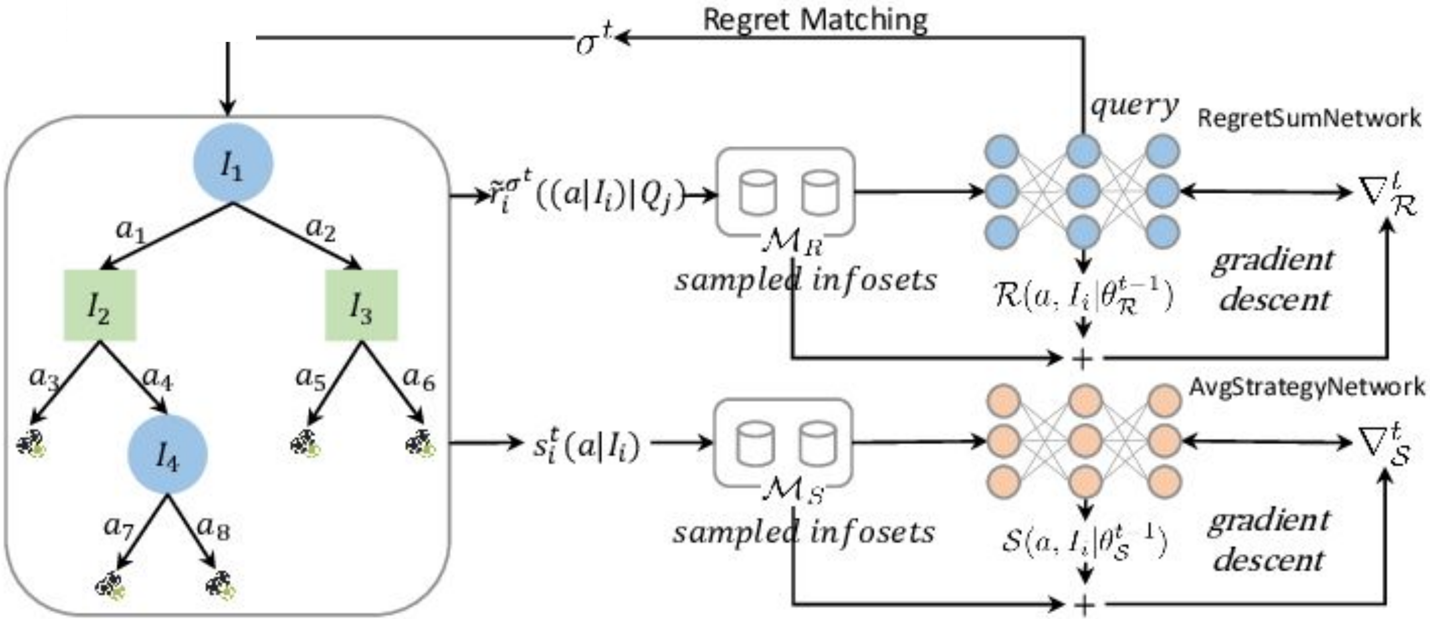
Double Neural CFR

- Proposes a double neural representation for imperfect information games
  - One neural network → cumulative regret
  - The other one → average strategy
- Make neural learning efficient by developing several novel techniques:
  - Robust sampling method
  - Mini-batch Monte Carlo CFR (MCCFR)
  - Monte Carlo CFR Plus (MCCFR+)

Tabular Methods



Double Neural CFR Method



Double Neural CFR: Results

- The algorithm:
    - Converges better than deep reinforcement learning counterparts
    - Matches the performance of tabular based algorithms
    - Has a strong generalization and compression ability

- The sampling method:
    - Lower variance than the outcome sampling
    - More memory efficient than the external sampling

Double Neural CFR: Problems

- May not be theoretically sound
- The authors consider only small games

Deep Counterfactual Regret Minimization 2019

CFR Problems

- Needs abstraction to deal with extremely large games
    - Abstraction algorithms may miss important details of the game as they are usually manual and domain-specific
    - Determining a good abstraction needs knowledge of the equilibrium of the game and we need abstraction to find the equilibrium (chicken-and-egg problem!)

Deep CFR

- Claim to be the first non-tabular variant of CFR
- Approximates the behavior of tabular CFR from partial game tree traversals
- Uses function approximation with deep neural networks to approximate the behavior of tabular CFR on the full, unabstracted game

Deep CFR: Method

- Approximate the behavior of CFR without calculating and accumulating regrets at each infoset
  - Generalizes across similar infosets using function approximation via deep neural networks

Deep CFR: Results

- Converges to an ε-Nash equilibrium in two-player zero-sum game
- Outperforms Neural Fictitious Self Play (NFSP) (the prior leading function approximation algorithm for imperfect-information games
- Competitive with domain-specific tabular abstraction techniques

Conclusion

Conclusion

- By introducing the iterative regret minimization larger games were solved
- Most algorithms use CFR to find Nash equilibrium in extensive-form games
- MCCFR was introduced to reduce the computation time of CFR
- CFR+ used exhaustive search and regret-matching+ to achieve high empirical results in poker games
- Some research use deep learning methods in combination with CFR

Works we haven’t covered

- Applications of the algorithms across other games and other multi-agent setting aside poker
- More generalization. Solving games with more than two players
- What are other solution methods aside from Nash equilibrium
- Games with imperfect recall
- Pruning methods

References

[1] Bowling, M., Burch, N., Johanson, M., & Tammelin, O. (2015). Heads-up limit hold'em poker is solved. *Science*, *347*(6218), 145-149.

[2] Zinkevich, M., Johanson, M., Bowling, M., & Piccione, C. (2007). Regret minimization in games with incomplete information. Advances in neural information processing systems, 20.

[3] Lanctot, M., Waugh, K., Zinkevich, M., & Bowling, M. (2009). Monte Carlo sampling for regret minimization in extensive games. Advances in neural information processing systems, 22.

[4] Schmid, M., Burch, N., Lanctot, M., Moravcik, M., Kadlec, R., & Bowling, M. (2019, July). Variance reduction in monte carlo counterfactual regret minimization (VR-MCCFR) for extensive form games using baselines. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 2157-2164).

[5] Tammelin, O. (2014). Solving large imperfect information games using CFR+. arXiv preprint arXiv:1407.5042.

[6] Li, H., Hu, K., Ge, Z., Jiang, T., Qi, Y., & Song, L. (2018). Double neural counterfactual regret minimization. arXiv preprint arXiv:1812.10607.

References

[7] Brown, N., Lerer, A., Gross, S., & Sandholm, T. (2019, May). Deep counterfactual regret minimization. In International conference on machine learning (pp. 793-802). PMLR.

[8] Sychrovsky, D., Sustr, M., Davoodi, E., Lanctot, M., & Schmid, M. (2023). Learning not to regret. arXiv preprint arXiv:2303.01074.

Questions