

A black dice cup with a red interior is shown on a green felt background. A red die is visible in the foreground. The title text is overlaid on a semi-transparent grid background.

Approximating Optimal Dudo Play with Fixed-Strategy Iteration Counterfactual Regret Minimization

Todd W. Neller
Steven Hnath
Gettysburg College

Outline

- Dudo Rules
- Regret and Counterfactual Regret
- Imperfect Recall Abstraction
- Counterfactual Regret Minimization (CFR)
Difficulties
- Fixed-Strategy Iteration Counterfactual Regret
Minimization (FSICFR)
- Results

Dudo







































- Bluffing dice game
- Origin generally believed to be 15th c. Inca
- Many variants/names (Liar's Dice, Perudo, Bluff, Call My Bluff, Cacho, Cachito)
- Internationally popular
 - BoardGameGeek.com rank 272/55074 (~top ½%), November 10th, 2011



Dudo Overview

- Bluffing dice game for 2+ players
 - Each player rolls 5 dice concealed under cup
 - Players make successively bolder claims about *all* dice rolled until player challenges
 - Loser of challenge loses dice
 - Last player with dice wins

Dudo Claims

	1		1		1		1		1		
1		2		2		2		2		2	
	3		3		3		3		3		
2		4		4		4		4		4	
	5		5		5		5		5		
3		6		6		6		6		6	
	7		7		7		7		7		

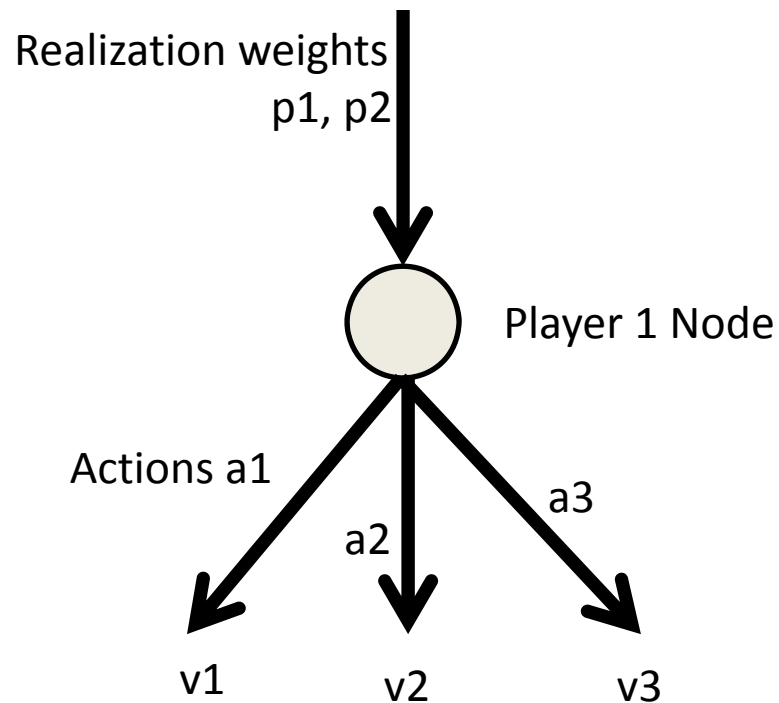
Dudo Rules

- Players each roll and privately view 5 dice concealed under a cup. 1s are wild.
- Players make successively greater claims until one challenges the previous claim with “Dudo!” (Sp. “I doubt it!”), all reveal dice, and:
 - More/less than claimed? Challenger/claimant loses dice according to difference.
 - Claim exactly correct? All others lose 1 die.
- The next round begins with challenge winner.
- The last player with dice wins.

Regret

- Rock, Paper, Scissors (a.k.a. Roshambo)
- +1 win, 0 tie, -1 loss
- Losing choice results in regret of $1/2$ for not choosing tie/win play.
- Hart & Mas-Colell Regret-Matching (2000):
 - Depart from current play with probabilities proportional to cumulative past regrets

Counterfactual Regret Example



- Input: realization weights
- Compute node strategy from normalized positive cumulative regret.
- Update avg. output strategy weighted by player realization weight.
- Recursively evaluate strategy to compute action values and node value.
- Compute counterfactual regret.
- Update cumulative regret weighted by opponent realization weight.

Counterfactual Regret Example

	p1	p2	
Realization Weights	0.5	0.25	
Player 1 Node:			
	a1	a2	a3
Cumulative Regret	20	-10	30
Positive Regret	20	0	30
Strategy	0.4	0	0.6
Cumulative Strategy +=	0.2	0	0.3
Player 1 Node Actions:	1	2	3
p1'	0.2	0	0.3
p2'	0.25	0.25	0.25
v1	40	-8	20
Node Value	28		
Action Regrets	12	-36	-8
Counterfactual Regrets	3	-9	-2
Old Cumulative Regret	20	-10	30
New Cumulative Regret	23	-19	28

- Input: realization weights
- Compute node strategy from normalized positive cumulative regret.
- Update avg. output strategy weighted by player realization weight.
- Recursively evaluate strategy to compute action values and node value.
- Compute counterfactual regret.
- Update cumulative regret weighted by opponent realization weight.

Dudo Information Sets

- Information set – all game states consistent with what you know
 - Dudo: Your dice, the number of opponent dice, the claim history
 - 2-player Dudo info sets:
294,021,177,291,188,232,192
- Imperfect recall abstraction
 - Remember up to the last m claims
 - $m = 3 \rightarrow 21,828,536$ abstract info. sets
 - Abstraction + CFR can lead to pathological behavior, but not apparently in this case.

Problems in applying CFR to Dudo

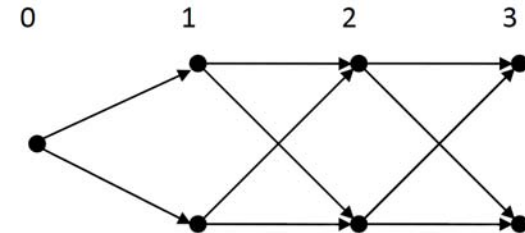
- Even with imperfect recall, the exponentially growing paths through possible states make application infeasible:

Dice	1	2	3	4	5
1	13	49	1551	97210	6179541
2	26	1507	99019	6310784	-
3	1509	98483	6211111	-	-
4	97611	6265326	-	-	-
5	6290658	-	-	-	-

(b) Time (ms) per single CFR training iteration.

Insight

- CFR w/ abstracted info. sets,
 - node visits grow exponentially with depth.
 - strategies and regrets change each node visit
- We have restructured the algorithm to make a single dynamic-programming style
 - forward pass computing strategies and accumulating realization weights, and
 - backward pass computing utilities and accumulating regrets.
- By fixing-strategies throughout the forward pass, we can use the same regret-matching approach, yet reduce exponential complexity to linear.



2-versus-2 Dice Training

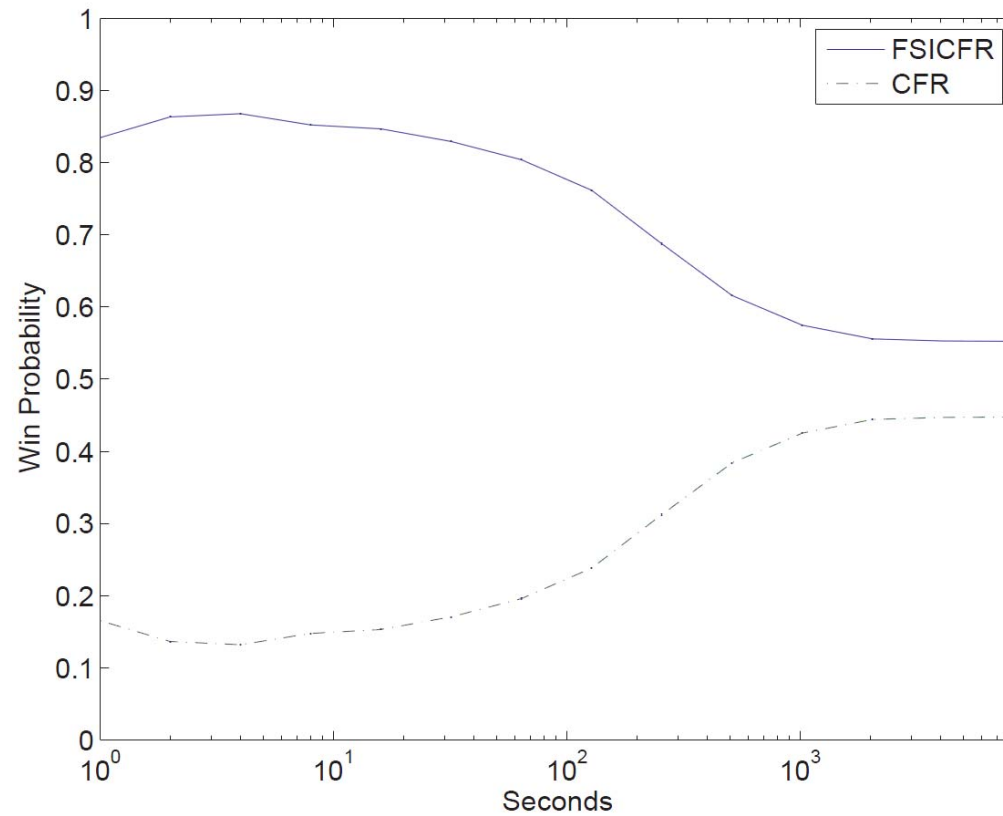


Fig. 3. FSICFR vs. CFR win rates during 2-vs.-2 dice training.

Time (ms) per Training Iteration

Dice	Opponent Dice				
	1	2	3	4	5
1	0.3	1.0	2.1	4.9	9.9
2	0.7	2.0	4.7	9.7	18.7
3	2.1	4.7	9.2	18.6	34.4
4	5.0	9.8	18.7	32.4	55.3
5	10.0	18.9	34.8	56.1	94.5

(a) Time (ms) per FSICFR training iteration averaged over 1000 training iterations.

Dice	Opponent Dice				
	1	2	3	4	5
1	13	49	1551	97210	6179541
2	26	1507	99019	6310784	-
3	1509	98483	6211111	-	-
4	97611	6265326	-	-	-
5	6290658	-	-	-	-

(b) Time (ms) per single CFR training iteration.

Fig. 4. Time (ms) per training iteration.

Varying Imperfect Recall

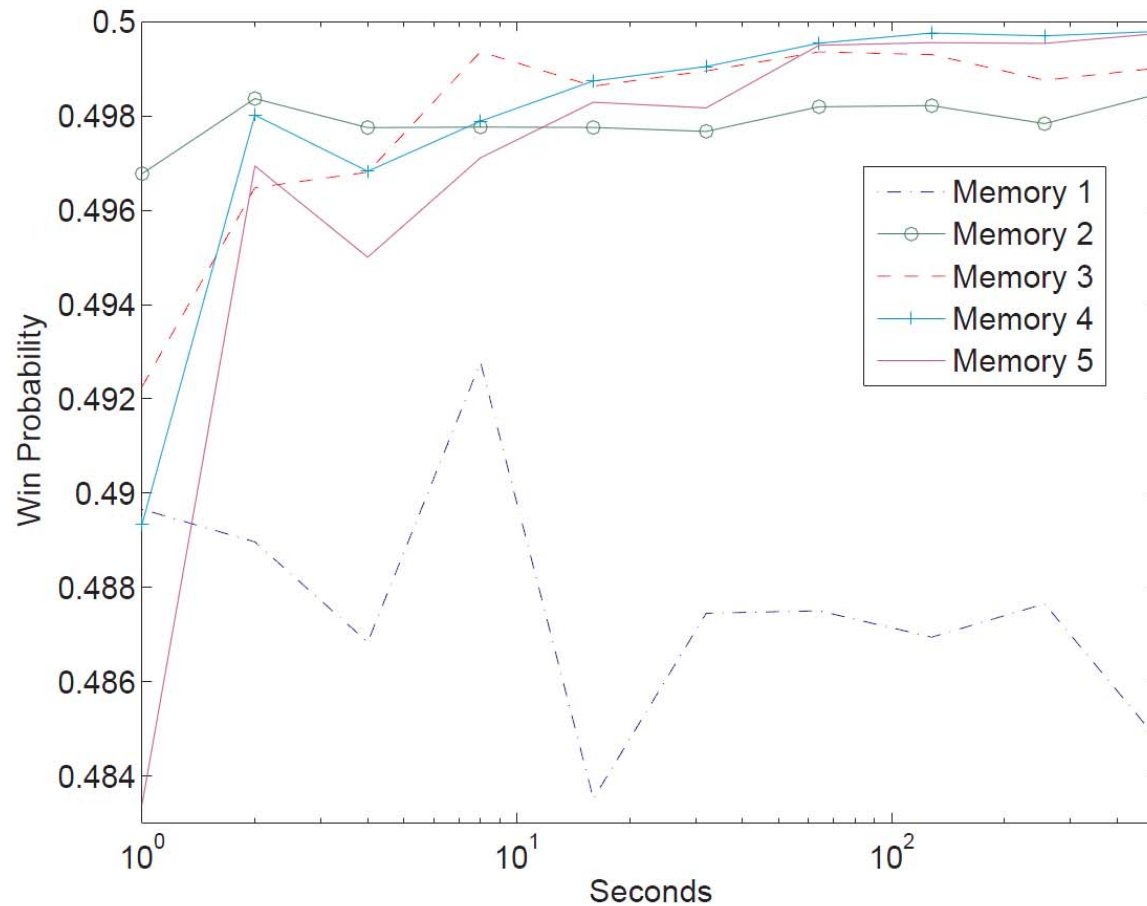


Fig. 5. 1-vs.-1 die FSICFR win rates vs. optimal strategy varying action recall imperfection.

Summary

- Imperfect recall abstraction brings Dudo information sets from 2.9×10^{20} to 2.2×10^7 .
- Even so, CFRs exponential growth through depths of claim sequences proves infeasible.
- FSICFR: Holding strategies fixed through dynamic-programming-style forward and backward pass allows efficient CFR training.
- Any extensive game tree with significant paths to the same (abstracted) info. sets benefits.