Stochastic Games from the viewpoint of Computational Complexity Theory

Kristoffer Arnsfelt Hansen - Aarhus University

Computational Complexity

Study of Computational Problems and Models of Efficient Computation

Typically **worst-case, polynomial** time, space, …

Examples:

- Given a 2-player zer … te optimal strategies.
- Given a finite stoch … mpute a NE for λ-discounted payoffs.
- Given a Linear Program, compute optimal solution. (?)
- Given a Boolean formula φ in variables $x_1, \ldots, x_\ell$, is there $x \in \{0,1\}^\ell$ such that $\varphi(x)$ is true? (**SAT** problem).
- Given existential first-order formula φ over $\mathbb{R}$ in variables $x_1, \ldots, x_\ell$, is there $x \in \mathbb{R}^\ell$ such that $\varphi(x)$ is true? (**ETR** problem).

Linear Programming as Computational Problem

Consider LP P: max $c^\intercal x$, s.t $Ax \leq b$
There might not be optimal solution, so what to compute?

Solution #1 – Maintain problem as **optimization problem**:
Given P, output either optimal x, *"infeasible"* or *"unbounded"*.

Solution #2 – Convert to **decision problem**:
Given P and k, is there x such that $c^\intercal x \geq k$ and $Ax \leq b$?

Solution #3 – Convert to **total search problem**:
Given P, output either optimal x, witness y of infeasibility (Farkas), or unbounded ray (x, z) witnessing unboundedness.

Comparing Problems : Reductions

Let P_1 and P_2 be computational problems. We consider P_1 to be **no harder** than P_2 if we can efficiently *reduce* P_1 to P_2.

Many-one reductions

- Instance mapping f maps instance I of P_1 to instance $f(I)$ of P_2.
- Solution mapping g maps solution y to $f(I)$ to solution $g(x)$ of I.

Turing reductions

Given a black box solving P_2 (for free), we can efficiently solve P_1.

Universal Problem: Completeness

Consider a class $\mathcal{C}$ of computational problems.

A computational problem P is *complete* for $\mathcal{C}$ if:

- $P \in \mathcal{C}$ (membership)
- Any $P' \in \mathcal{C}$ reduces to P (hardness)

Amazing fact: Virtually all ”natural” computational problems are complete for one of very few classes of computational problems.

Example: Equivalence of LP and matrix games

We can compute optimal strategies of matrix game by a simple LP.

Dantzig (1951) Given LP max $c^\intercal x$, s.t $Ax \leq b$, construct symmetric matrix game G as

$$\mathrm{G} = \begin{pmatrix} 0 & A & -b \\ -A^\intercal & 0 & c \\ b^\intercal & -c^\intercal & 0 \end{pmatrix}$$

Theorem: Let (y, x, t) be optimal strategy in G. If $t > 0$ then x/t and y/t are optimal solutions to primal and dual LP.

Widely and *incorrectly* cited to show equivalence of solving LP and matrix games.

Adler (2012) *Strongly polynomial time* reduction of solving LP to solving matrix games.

Also, **Brooks & Reny (2021)**

Matrix games: FP and P

Consider matrix game G

$$\begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

Search problem:

Given G, Compute optimal strategies (x, y)

FP

Decision problem (example):

Given G and $p > 0$, Is there optimal strategies (x, y), where no pure strategy is played with probability more than p?

P

Bimatrix games: PPAD and NP

Consider bimatrix game G

$$\begin{pmatrix} 4,4 & 3,5 \\ 5,3 & 1,1 \end{pmatrix}$$

Search problem:

Given G, Compute NE (x, y)

PPAD complete

Daskalakis, Goldberg, and Papadimitriou, 2006

Chen and Deng, 2006

Decision problem (example):

Given G and $p > 0$, Is there NE (x, y), where no pure strategy is played with probability more than p?

NP-complete

Gilboa and Zemel, 1989.

Conitzer and Sandholm, 2008

Nornal form games: FIXP and $\exists\mathbb{R}$

Consider 3-player game G

	1	2
1	$(-4,2,2)$	$(-2,1,1)$
2	$(-2,1,1)$	$(0,0,0)$

1

	1	2
1	$(0,0,0)$	$(-2,1,1)$
2	$(-2,1,1)$	$(-6,3,3)$

2

Search problem:

Given G, Compute NE (x, y, z)

FIXP complete

Etessami and Yannakakis, 2010.

Decision problem (example):

Given G and $p > 0$, Is there NE (x, y, z), where no pure strategy is played with probability more than p?

$\exists\mathbb{R}$ -complete

Schaefer and Štefankovič , 2017

Garg, Mehta, Vazirani, and Yazdanbod, 2018

Bilò and Mavronicolas, 2021

Berthelsen and Hansen, 2022

PPAD vs FIXP

PPAD

- Total search problems in NP
- Finding ε-almost Brouwer fixed points of continuous function F
- Finding ε-NE in normal form games
- Finding exact NE in bimatrix games
- Combinatorial hardness

Papadimitriou, 1994

Daskalakis, Goldberg, and Papadimitriou, 2006

Chen and Deng, 2006

FIXP

- Total search problems in $\exists\mathbb{R}$
- Finding (exact) Brouwer fixed points of of continuous function F
- Finding (exact) NE in normal form games
- Combinatorial and Numerical hardness

Etessami-Yannakakis, 2007

Discounted Stochastic Games - hardness

Problem: Given multiplayer stochastic game G and discount λ, compute λ-discounted NE.

Normal form games are special cases, and hence:

PPAD-hard for 2-player games and FIXP-hard for $\geq$3-player games.

Etessami and Yannakakis, 2007

Let G be 2-player zero-sum game, λ discount, s a state, and k rational number.

Deciding if $\mathrm{val}_\lambda(s) \geq k$ is SQRT-SUM-hard.

SQRT-SUM: Given positive integers $d_1, \ldots, d_n, k$, decide whether $\sum_{i=1}^{n} \sqrt{d_i} \leq k$.

Discounted Stochastic Games - membership

Problem: Given multiplayer stochastic game G and discount λ, compute λ-discounted NE.

Etessami and Yannakakis, 2007

For 2-player zero-sum game G problem is in FIXP.

Computing strategy profile ε-close to NE is in PPAD.

Filos-Ratsikas, Hansen, Høgh, and Hollender, 2021

General problem is in FIXP (and is hence FIXP-complete).

Proof of existence of λ-discounted NE

Takahashi, 1964 & Fink, 1964 Proof using Kakutani's FPT.

$$u_i^{s,\lambda,v}(a) = \lambda u_i(s,a) + (1-\lambda)\sum_{s'} q(s'|s,a)v_i(s')$$

$$D = \left([-M,M]^S\right)^n \times (\Delta(A_1)^S \times \cdots \times \Delta(A_n)^S)$$

$$F : D \rightrightarrows D$$

$$F(v,x) = (G(v,x), H(v,x))$$

$$G(v,x)_{i,s} = \max_{y(s)_i} u_i^{s,\lambda,v}(y(s)_i\ ; x(s)_{-i})$$

$$H(v,x)_{i,s} = \operatorname*{argmax}_{y(s)_i} u_i^{s,\lambda,v}(y(s)_i\ ; x(s)_{-i})$$

Proof by Brouwer's FPT

$$u_i^{s,\lambda,v}(a) = \lambda u_i(s,a) + (1-\lambda)\sum_{s'} q(s'|s,a)v_i(s')$$

$$\widetilde{D} = \left([-M,M]^S\right)^n \times \left(\Delta(A_1)^S \times \cdots \times \Delta(A_n)^S\right)^2$$

$$\tilde{F}: \widetilde{D} \to \widetilde{D}$$

$$\tilde{F}(v,x,y) = (\tilde{G}(v,x,y), y, \widetilde{H}(v,x,y))$$

$$\tilde{G}(v,x,y)_{i,s} = u_i^{s,\lambda,v}(y(s)_i\,; x(s)_{-i})$$

Choose $\widetilde{H}: \widetilde{D} \to \left(\Delta(A_1)^S \times \cdots \times \Delta(A_n)^S\right)$ to satisfy that $\widetilde{H}(v,x,y) = y$ implies $y \in H(v,x)$, i.e., y are best replies to x in game given by u.

Further results on zero-sum games

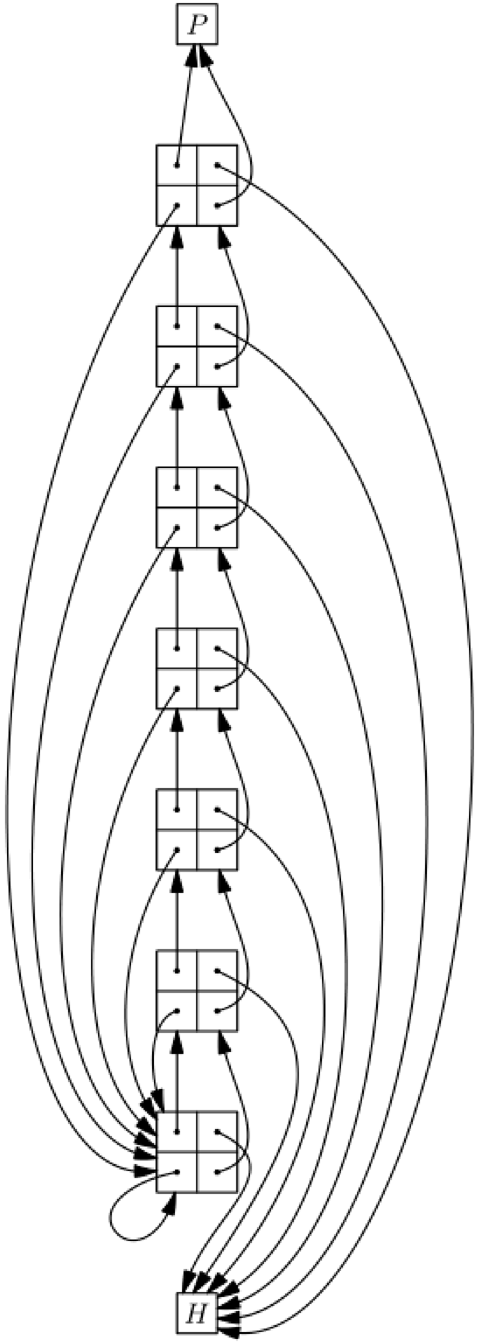
Condon, 1992 "Simple Stochastic Games" : Zero-sum, perfect information stochastic games with non-negative terminal rewards.

Juba, 2005 Problem of computing optimal strategies is in PPAD.

Theorem Given zero-sum stochastic games G and $\varepsilon > 0$, computing additive ε-approximation of v_∞ is in FIXP

Proof: R.A.G. implies that $\|v_\lambda - v_\infty\|_\infty \leq \varepsilon$ for $\lambda < \varepsilon^{2^{|G|^{O(1)}}}$

We can compute such small λ by repeated squaring efficiently inside FIXP function (the Brouwer function).



Purgatory Game

Hansen, Koucký, Miltersen, 2009

n	v_n
1	0.01347
10	0.03542
100	0.06879
1000	0.10207
10000	0.13396
100000	0.16461
1000000	0.19415
10000000	0.22263
100000000	0.24828
$> 10^{128}$	0.99

Hansen, Ibsen-Jensen, Miltersen, 2009

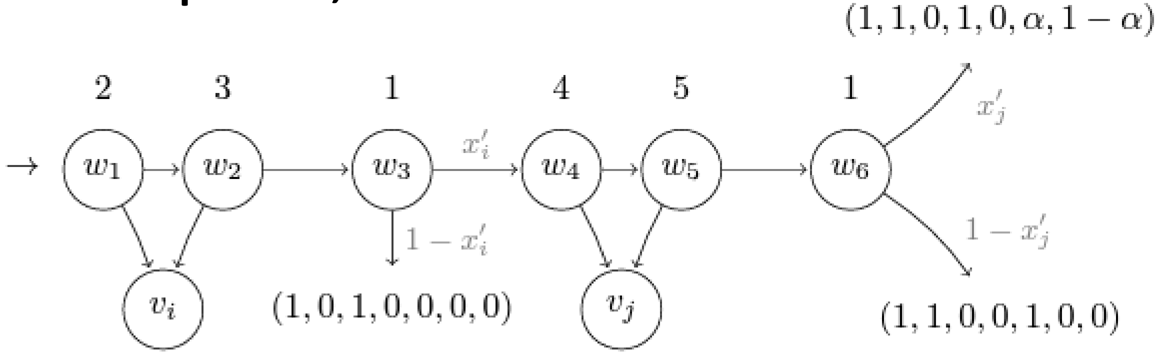
Doubly-exponential lower bound on n (as a function of number of states) to ensure v_n is non-trivial approximation of v_∞.

Multi-Player Perfect-Information Games

Let G be a perfect information stochastic game with terminal rewards.

Ummels and Wojtczak, 2001 Study of various decision problems, giving NP-hardness/completeness, PSPACE-hardness/completeness, SQRT-SUM-hardness, undecidability,…

Hansen and Sølvsten, 2020 Deciding existence of stationary NE in m-player games is $\exists\mathbb{R}$-complete, for $m \geq 7$.



Back to zero-sum stochastic games

Gillette, 1957 The Big Match

Blackwell and Ferguson, 1968
Existence of ε-optimal strategy for Player 1

	1	-1
C	1	-1
A	-1*	1*

At stage t, let $k_t = \sum_{j=1}^{t-1} r_j$ be the sum of rewards in previous stages.
Play the absorbing action with (conditional) probability

$$\frac{1}{(N + k_t)^2}$$

where $N = N_\varepsilon$ is sufficiently large.

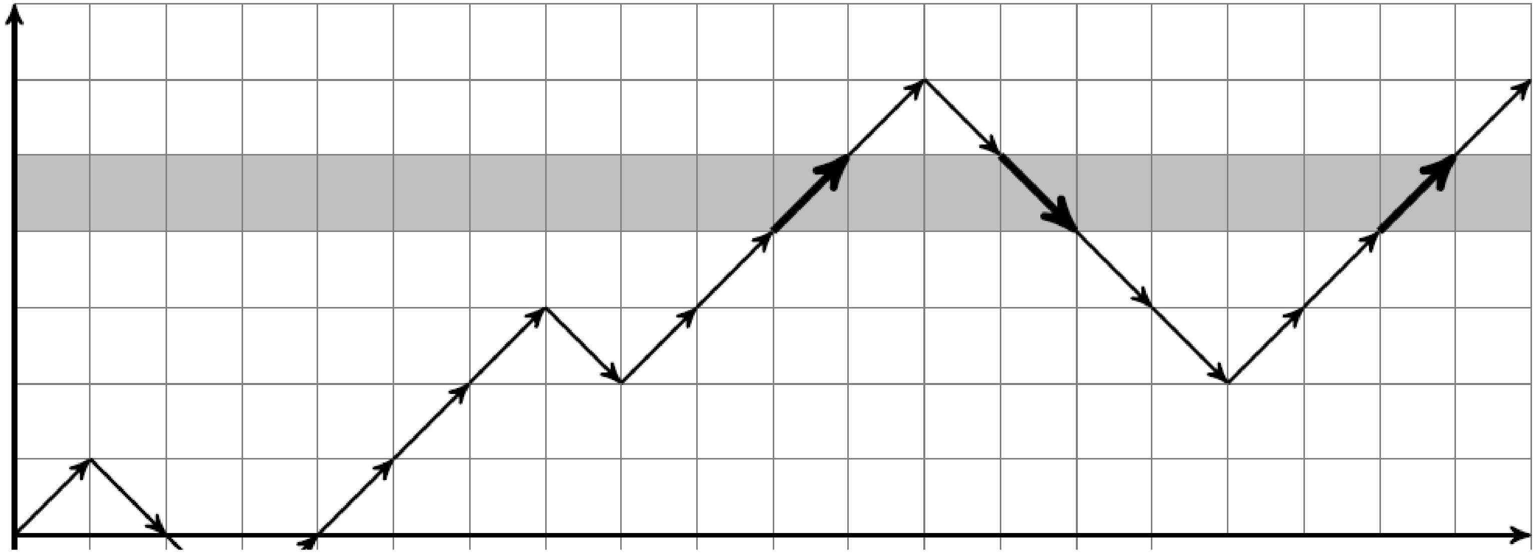
Blackwell and Ferguson, 1968

Alternative strategy:
In stage t play absorbing action with probability

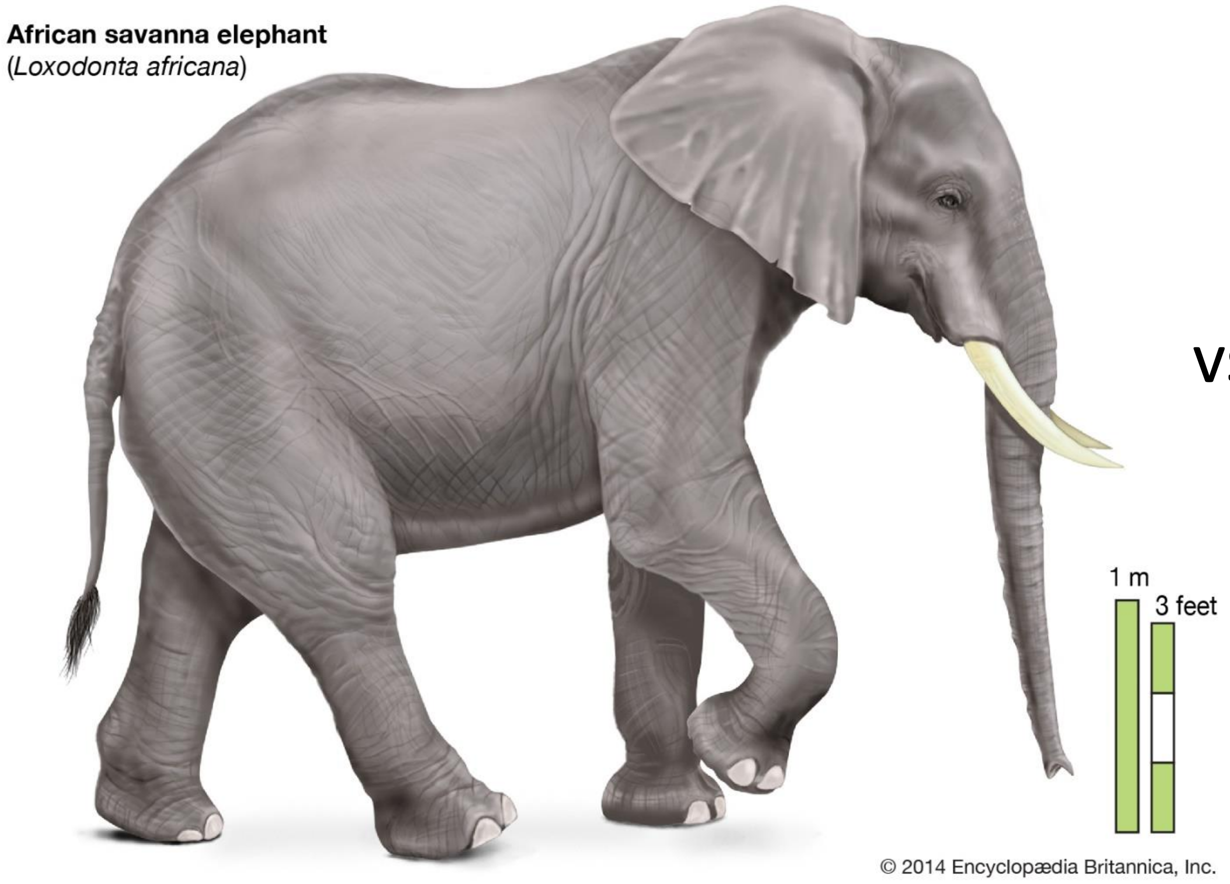
(conditioned on reaching stage t without absorbing)

$\epsilon^2(1-\epsilon)^{\max(k_t,0)}$, $k_t = \sum_{j=1}^{t-1} r_j$

where r_i is the reward in stage i.



Memory requirement



vs

Inspiration: Probabilistic Approximate Counting

Morris, 1978 Flajolet, 1985

Maintain probabilistic counter

Count to n with constant relative accuracy using only $\log_2 \log_2 n + O(1)$ bits.

Programming Techniques

S.L. Graham, R.L. Rivest Editors

Counting Large Numbers of Events in Small Registers

Robert Morris
Bell Laboratories, Murray Hill, N.J.

It is possible to use a small counter to keep approximate counts of large numbers. The resulting expected error can be rather precisely controlled. An example is given in which 8-bit counters (bytes) are used to keep track of as many as 130,000 events with a relative error which is substantially independent of the number *n* of events. This relative error can be expected to be 24 percent or less 95 percent of the time (i.e. $\sigma = n/8$). The techniques could be used to advantage in multichannel counting hardware or software used for the monitoring of experiments or processes.

Key Words and Phrases: counting
CR Categories: 5.11

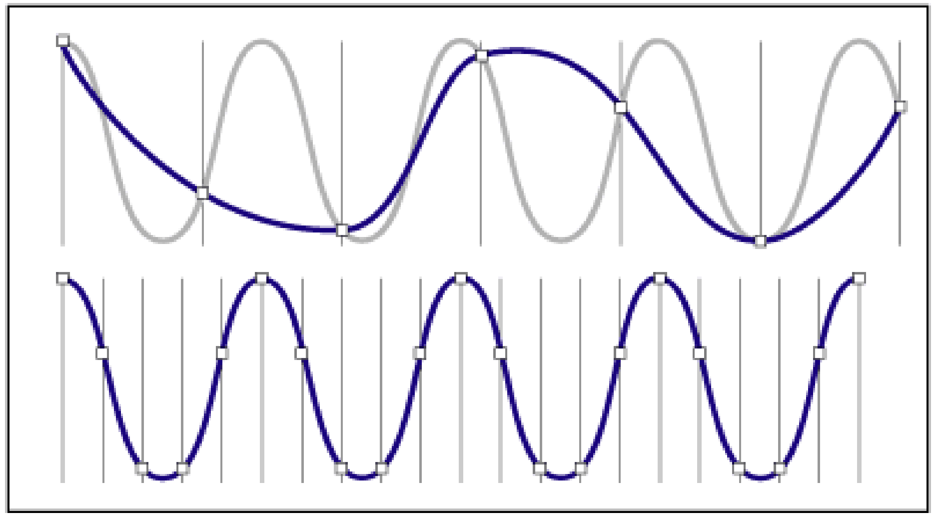
Small-space strategies

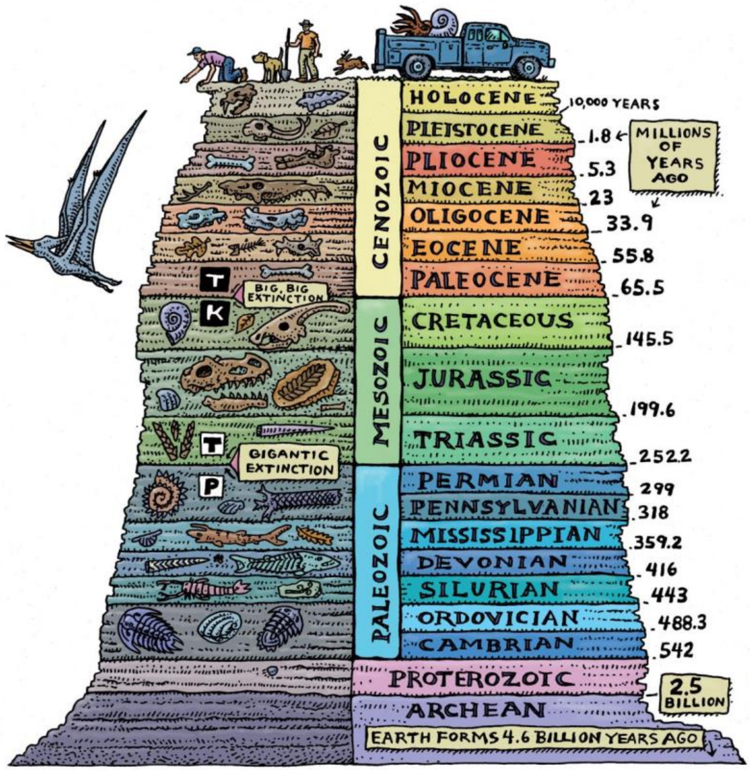
Hansen, Ibsen-Jensen, and Koucký, 2016

Theorem For all $\varepsilon > 0$, there is a **liminf** ε-optimal strategy in the Big Match that for any $\delta > 0$ with probability $\geq 1 - \delta$ the uses $(\log n)^{O(1)}$ memory states in round n. (i.e. $\mathrm{O}(\log \log n)$ bits of memory).

Theorem For all $\varepsilon > 0$, and **any** non-decreasing unbounded function s, there is a limiting average **limsup** ε-optimal strategy in the Big Match that for any $\delta > 0$ with probability $\geq 1 - \delta$ the uses $O(s(n))$ memory states in round n.

Alternative ideas: Sampling and Epochs





Finite Memory Strategies

Hansen, Ibsen-Jensen, and Neyman, 2018

For all $\varepsilon > 0$, there is a clock-dependent strategy σ using 2 states of memory that is ε-optimal for Player 1.