Simple Stochastic ω-Regular Games

Krishnendu Chatterjee

PAuL 2006

Stochastic Games

- Games played on game graphs with stochastic transitions.

- Stochastic games [Sha53]
  - Framework to model natural interaction between components and agents.
    - e.g., controller vs. system.

Games

- Where:
  - Arena: Game graphs.
- What for:
  - Objectives - ω-regular.
- How:
  - Strategies.

Key Issues

- Characterize the structure of winning.
- Algorithms and complexity.
- Strategy classification: simplest class of strategy that suffices for winning.

Turn-based (simple) Stochastic Games

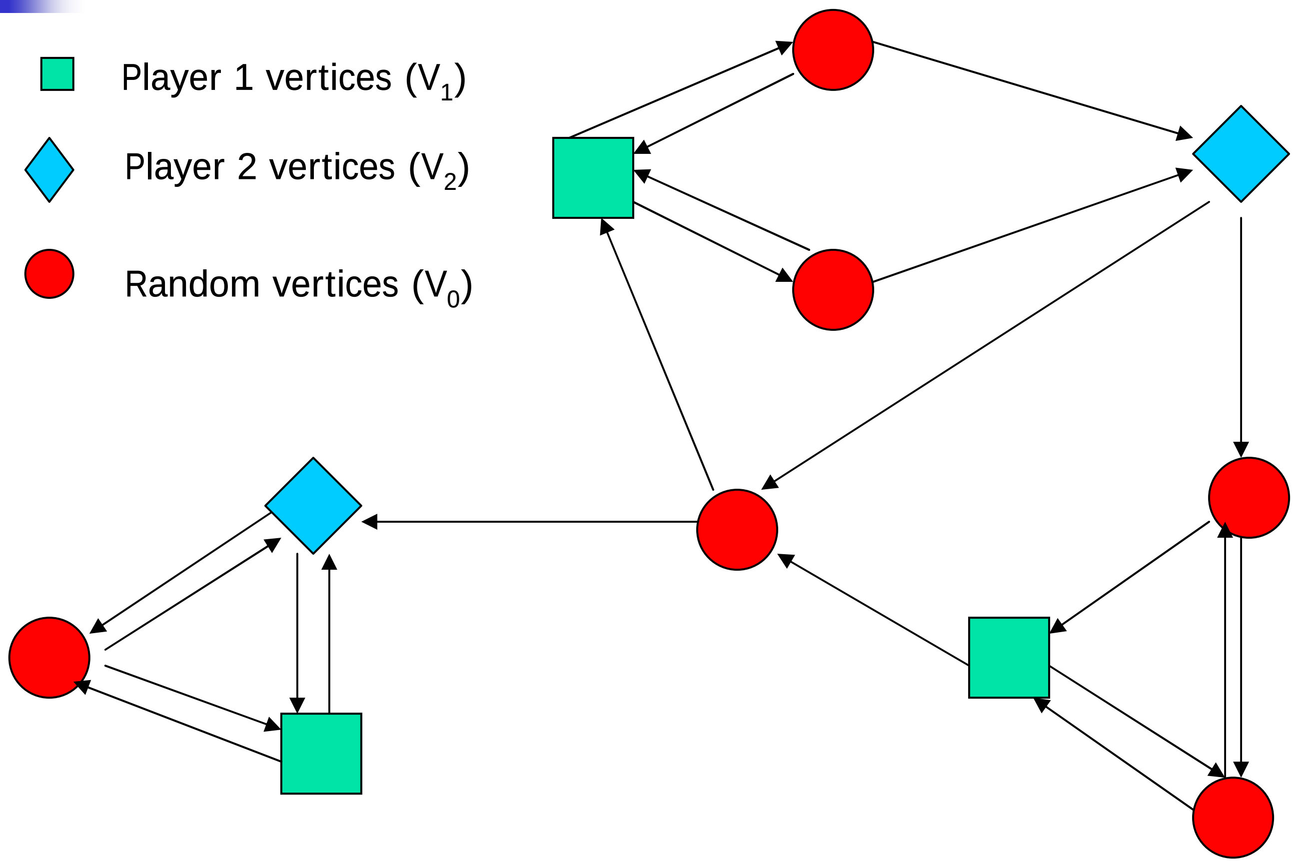
Turn-based Stochastic Games

- A turn-based stochastic game is defined as
    - $G=((V,E),(V_1,V_2,V_0))$, where
        - (V,E) is a graph with a finite vertex set V.
        - (V_1,V_2,V_0) is a partition of V.
- V_1 player 1 makes moves.
- V_2 player 2 makes moves.
- V_0 randomly chooses successors according to a probability distribution δ.

Game Played by Moving Token

- Token placed on an initial vertex.
- If current vertex is
  - Player 1 vertex then player 1 chooses successor.
  - Player 2 vertex then player 2 chooses successor.
  - Player random vertex proceed to successors according to the probability distribution δ.
- Generates infinite sequence of vertices.

A Turn-based Stochastic Game



Sub-classes

- Markov decision processes (MDPs)
  - $V_1 = \emptyset$ or $V_2 = \emptyset$
  - 1 player and randomness (refer as 1 ½ player games).
  - Games against nature or randomness.

- Turn-based deterministic games
  - $V_0 = \emptyset$
  - 2 player and no randomness (refer as 2 player games).

- Turn-based stochastic games
  - 2 player and randomness (refer as 2 ½ player games).

Applications

- Markov decision processes
  - Control in presence of uncertainty [Put94, FV 97].
  - AI planning.
- 2 player games
  - Controller synthesis [RW 87, PR89].
- 2 ½ player games
  - Controller synthesis in presence of uncertainty [BGLBC 04].
  - Robust version of MDPs.

Objectives

Objectives

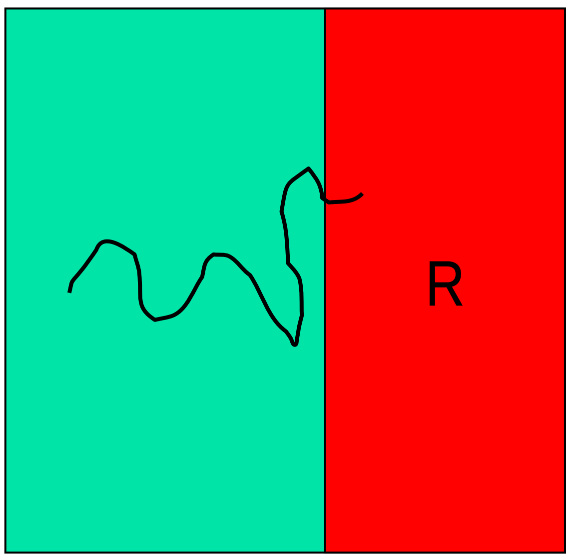
- Plays: infinite sequence of vertices.
- Objectives: subset of plays, $\Psi_1 \subseteq V^\omega$.
- Play is winning for player 1 if it is in Ψ_1.
- Zero-sum game: $\Psi_2 = V^\omega \setminus \Psi_1$.

Reachability and Safety

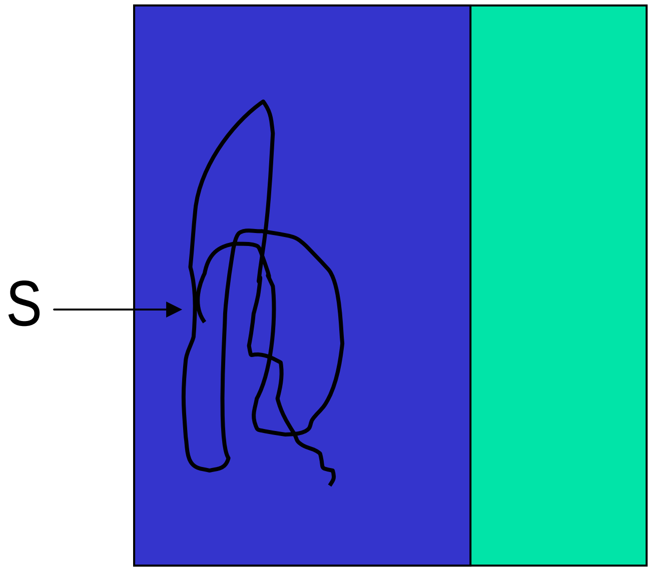
- Let $R \subseteq V$ set of target vertices. Reachability objective requires to visit the set R of vertices.

- Let $S \subseteq V$ set of safe vertices. Safety objective requires never to visit any vertex outside S.

Reachability and Safety



Reachability



Safety

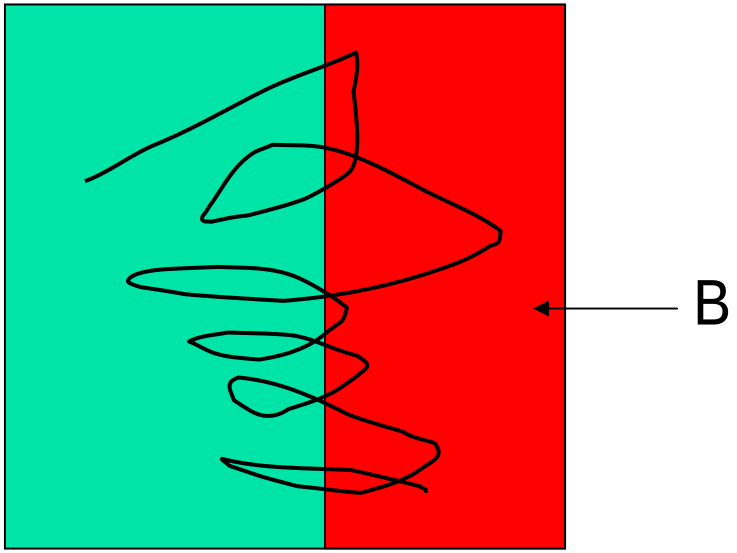
Buechi Objective

- Let $B \subseteq V$ a set of Buechi vertices.

  Buechi objective requires that the set B is visited infinitely often.

- Correspond to liveness objectives.

Buechi



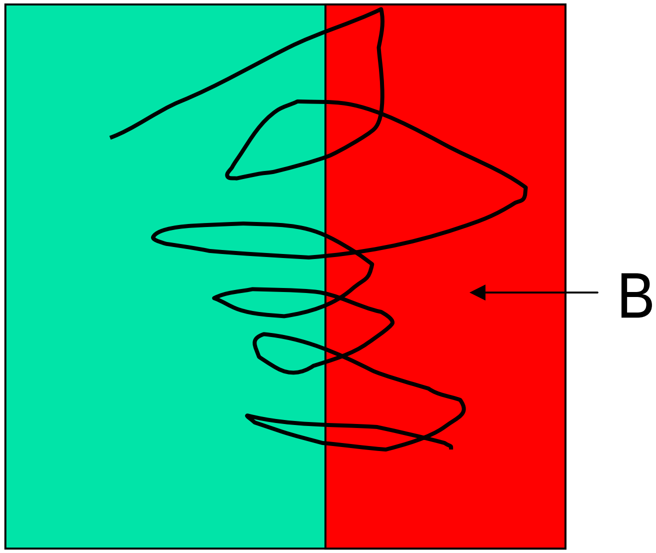
Buchi

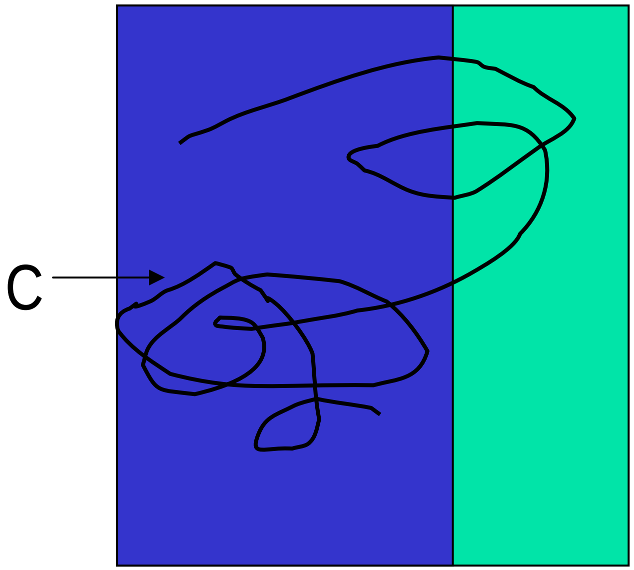
coBuechi Objective

- Let $C \subseteq V$ a set of coBuechi vertices. coBuechi objectives requires that vertices outside C visited finitely often.

- Dual of Buechi objective.

Buchi and coBuchi





Buchi

coBuchi

Rabin and Streett Objectives

- Let $\{(E_1,F_1), (E_2,F_2),\ldots, (E_d,F_d)\}$ set of vertex set pairs.
  - Rabin: requires there is a pair (E_j,F_j) such that E_j finitely often and F_j infinitely often.
    - $\exists$ j. (E_j finitely $\wedge$ F_j infinitely)
  - Streett: requires for every pair (E_j,F_j) if F_j infinitely often then E_j infinitely often.
    - $\forall$ j. (F_j finitely $\vee$ E_j infinitely) = $\forall$ j. (F_j infinitely $\rightarrow$ E_j infinitely)
    - Conjunction of Fairness conditions (strong fairness).

Rabin-chain Objectives

- Rabin-chain: both a Rabin-Streett, complementation closed subset of Rabin.
- Rabin, Streett and Rabin-chain canonical forms to express ω–regular objectives
  - Express liveness, fairness, etc.

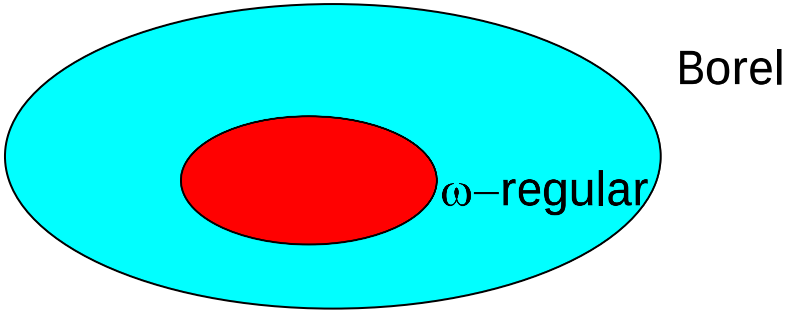
Mueller Objectives

- Let $M \subseteq 2^V \setminus \emptyset$ be a subset of vertices. The Mueller objective requires the set of vertices appearing infintely often is in M.

- Remarks
  - Buechi and coBuechi objectives are simple classes of Rabin-chain objectives.
  - Rabin and Streett objectives are special classes of Mueller objectives.
  - Mueller objectives can be reduced to Rabin-chain objectives.

Objectives

- ω-regular: ∪, ° , *, ω.
  - Safety, Reachability, Liveness, etc.
  - Mueller, Rabin, Streett, Rabin-chain express all ω-regular objectives.



Strategies

Strategies

- Given a finite sequence of vertices, (that represents the history of play) a strategy σ for player 1 is a probability distribution over the set of successor.

  - $\sigma : V^* \cdot V_1 \to \text{Dist}(V)$
  - Σ : set of player 1 strategies
  - Π : set of player 2 strategies.

Special Classes of Strategies

- Finite memory strategies: finite memory encoded as a finite automaton.
- Memoryless (stationary) strategies: Strategy is independent of the history of the play and depends on the current vertex.
  - $\sigma: V_1 \rightarrow Dist(V)$
- Pure strategies: chooses a successor rather than a probability distribution.
- Pure-memoryless: both pure and memoryless (simplest class).
  - $\sigma: V_1 \rightarrow V$

Values

- Given objectives Ψ_1 and $\Psi_2 = V^{\omega} \setminus \Psi_1$ the value for the players are
  - $val_1(\Psi_1)(v) = \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} Pr_v^{\sigma,\pi}(\Psi_1)$.
  - $val_2(\Psi_2)(v) = \sup_{\pi \in \Pi} \inf_{\sigma \in \Sigma} Pr_v^{\sigma,\pi}(\Psi_2)$.

Determinacy

- Determinacy:
  - $val_1(\Psi_1)(v) + val_2(\Psi_2)(v) = 1$.

- Determinacy means
  - sup inf = inf sup.
  - von Neumann's minmax theorem in matrix games.

Optimal Strategies

- A strategy σ is optimal for objective Ψ_1 if
  - $\mathrm{val}_1(\Psi_1)(v) = \inf_\pi \Pr_v^{\sigma,\pi} (\Psi_1)$.
- Analogous definition for player 2.

Results on Stochastic Games

MDPs

- Pure memoryless optimal strategy exists for Rabin objectives [CdAH04].

- Randomized memoryless or pure finite-memory optimal strategy exists for Mueller objectives [CdAH04].

- Reachability objectives with target T can be solved by the following linear program.
  - min $\sum_{v \in V} x_v$ subject to
    - $x_v \geq x_w \qquad v \in V_1, (v,w) \in E$
    - $x_v = \sum_{(v,w) \in E} x_w \, \delta(v,w) \qquad v \in V_0$
    - $x_v = 1 \qquad v \in T$

2 player Games

- 2 player games.
  - Determinacy (sup inf = inf sup) theorem for Borel objectives. [Mar75]
  - Values either 0 or 1.
  - Pure finite-memory optimal strategy exists for all Mueller objectives [GH82].
  - Pure memoryless optimal strategy exists for Rabin objectives. [EJ88]
    - NP-complete for Rabin and coNP-complete for Streett.
    - NP $\cap$ coNP for Rabin-chain objectives.

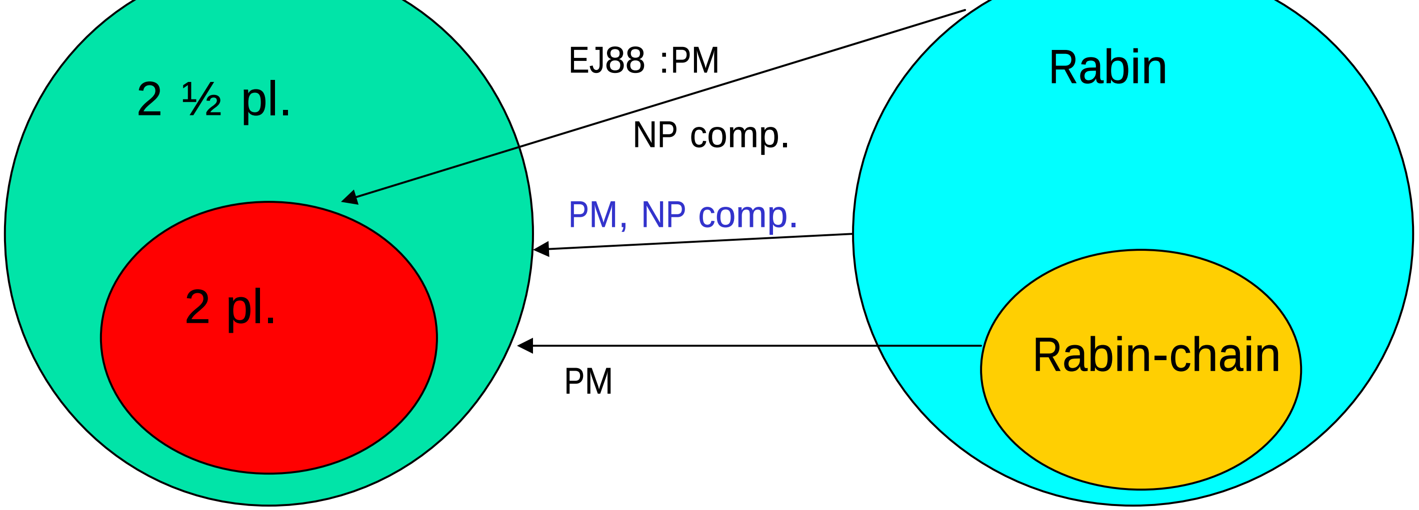
2 ½ player Games

- 2 ½ player games
  - Borel objectives: determinacy [Mar 98].
  - Reachability objectives: [Con92]
    - Pure memoryless optimal strategy exists.
    - Decided in NP $\cap$ coNP.
  - Rabin-chain objectives: [MM 03, CJH 04, Zie 04]
    - Pure memoryless optimal strategy exists.
    - Decided in NP $\cap$ coNP.
  - Rabin and Streett objectives [CdAH 05]
    - Pure memoryless optimal strategy exists for Rabin objectives.
    - NP-complete for Rabin objectives and coNP-complete for Streett objectives.

2 ½ player Games

- 2 ½ player games
  - Borel objectives: determinacy [Mar 98].
  - Reachability objectives: [Con92]
    - Pure memoryless optimal strategy exists.
    - Decided in NP $\cap$ coNP.
  - Rabin-chain objectives: [MM 03, CJH 04, Zie 04]
    - Pure memoryless optimal strategy exists.
    - Decided in NP $\cap$ coNP.
  - Rabin and Streett objectives [CdAH 05]
    - Pure memoryless optimal strategy exists for Rabin objectives.
    - NP-complete for Rabin objectives and coNP-complete for Streett objectives.

This Talk



Focus of This Talk

- Two basic methods for stochastic ω-regular games
  - Rabin objectives pure memoryless optimal strategy exists.
  - Optimal strategy construction for Mueller objectives.
  - Application of the methods.

Basic Methods

- Method 1: Reduction to two player games for qualitative analysis.
- Method 2: Construction of optimal strategies from qualitative winning strategies in sub-games.

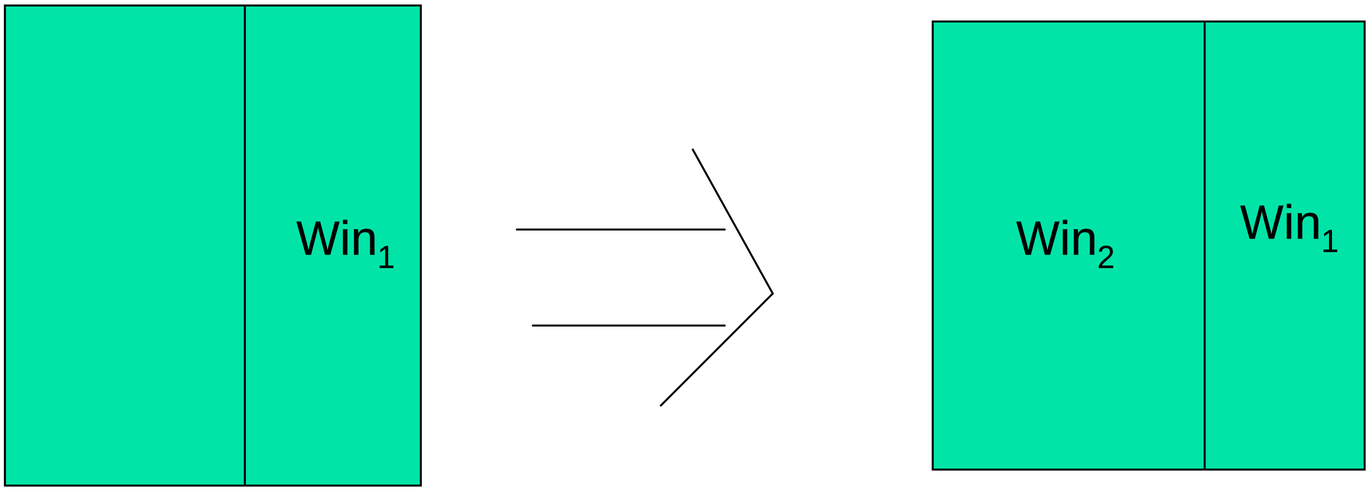
Method 1

Qualitative Analysis

- Qualitative analysis:
  - Analysis of value 1 and related optimal strategy.
  - W_1 ={ v : $val_1(\Psi_1)(v)$ =1 }, almost-sure winning states.

- Almost-sure winning strategy σ if
  - $\inf_{\pi} Pr_v^{\sigma,\pi}(\Psi_1)=1$

Reduction



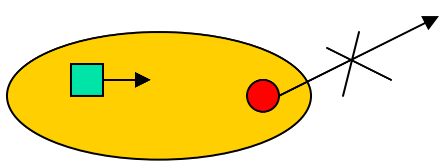
2 ½ player game

2 player game

MDPs and End-component

- A set of vertex C is an end-component if
  - C is strongly connected;
  - for all vertex $v \in V_0 \cap C$, if $(v,w) \in E$ then $w \in C$.

- Generalizes the notion of SCCs in graph and closed recurrent classes in Markov chains.



End-component

- [CY 95, deAl97] If U is not an end-component, then the probability that U is the set of vertices visited infinitely often is 0.

- Hence, with probability 1 the set of vertices visited infinitely often is a collection of end-components.

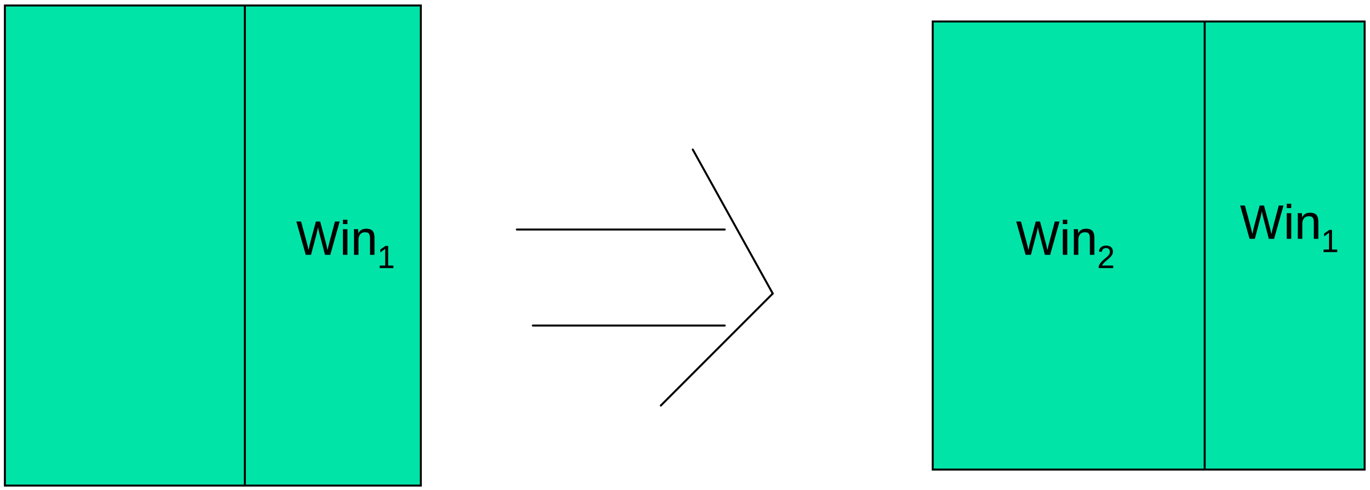
Rabin and Streett Objectives

- Let $\{(E_1,F_1), (E_2,F_2),\ldots, (E_d,F_d)\}$ set of vertex set pairs.
  - Rabin: requires there is a pair (E_j,F_j) such that E_j finitely often and F_j infinitely often.
  - Streett: requires for every pair (E_j,F_j) if F_j infinitely often then E_j infinitely often.

Winning End-component

- An end-component C is winning for player 1 if there is a pair j such that
  - $C \cap E_j = \emptyset$.
  - $C \cap F_j \neq \emptyset$.
- Otherwise winning for player 2.

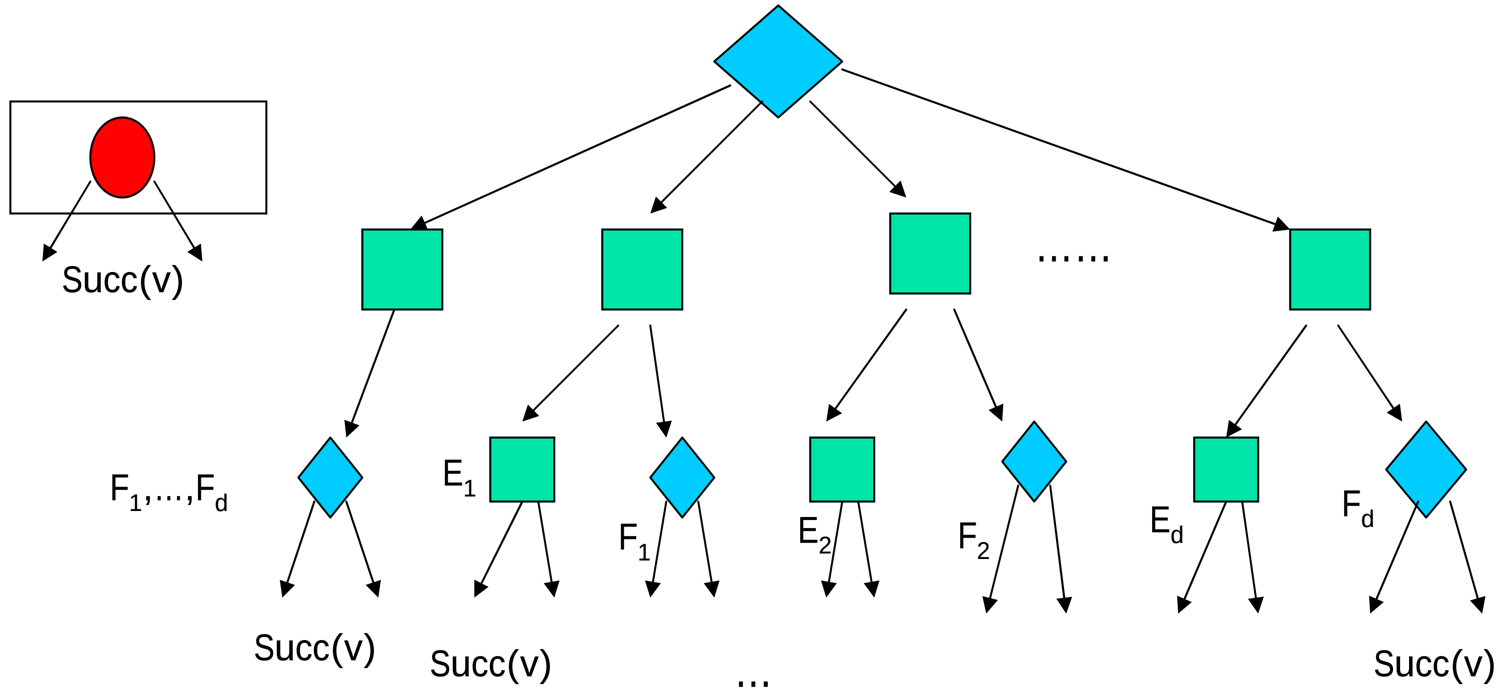
Reduction



2 ½ player game

2 player game

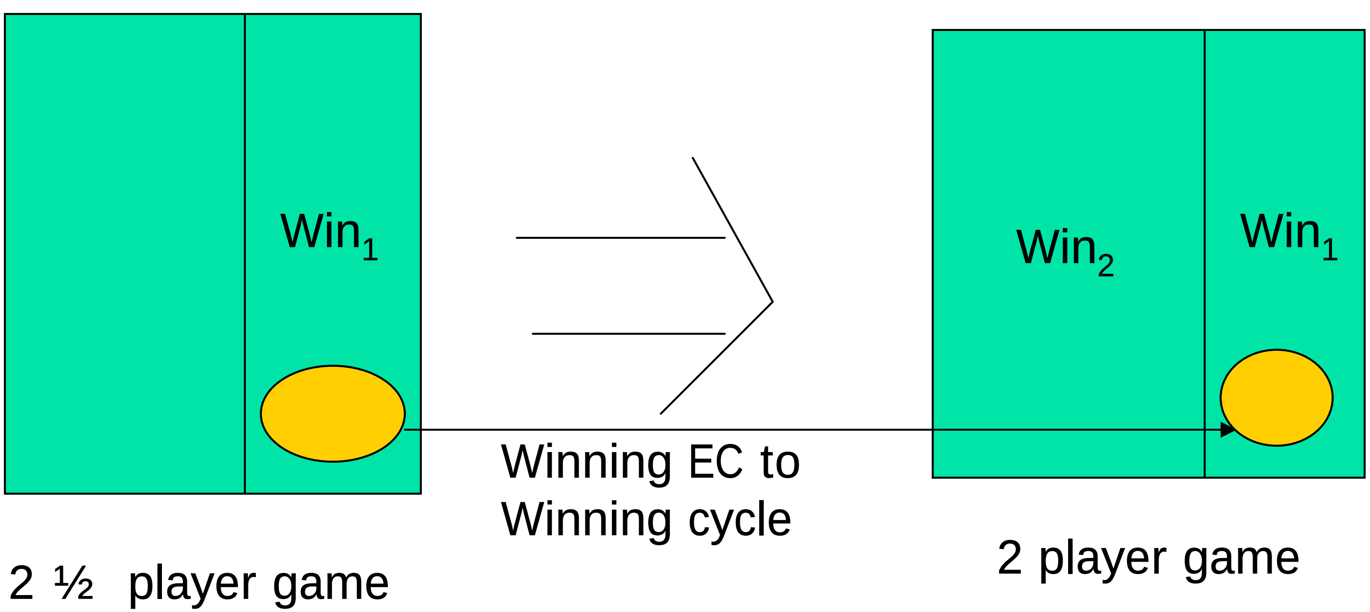
Reduction Gadget: Random Vertices



Proof Idea

- Get a pure memoryless optimal strategy from the 2-player game.
- Argue that strategy is almost-sure winning for player 1 in Win_1.

Reduction



Win$_1$
Win$_2$
Win$_1$
Winning EC to
Winning cycle
2 ½ player game
2 player game

Proof Idea

- Every end-component in Win_1 is winning for player 1.
- Every end-component in $V \setminus Win_1$ is winning for player 2.
  - Similar argument, constructing a winning cycle for player 1 from a winning end-component.

Almost-sure Winning

- In Win_1 it visits with probability 1 the end-components in Win_1.
- In $Win_2 = V \setminus Win_1$ visits with positive probability end-components in Win_2.
- Win_1 is the almost-sure winning set.

Qualitative Analysis

- Theorem: Pure memoryless almost-sure winning strategies exist in 2 ½ player Rabin games.

- Algorithms: Qualitative analysis for 2 ½ player games can be solved by algorithms for 2 player games.

Method 2

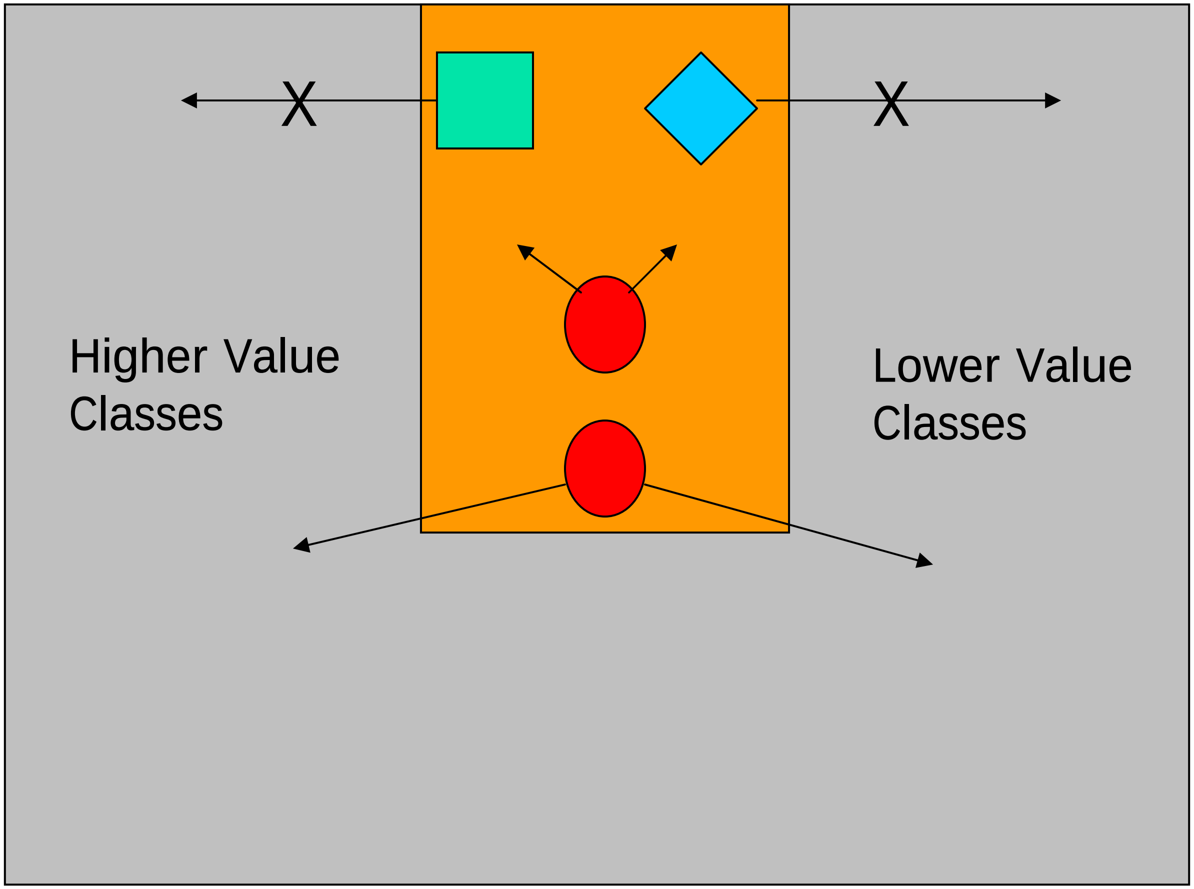
Value Class

- Value class is the set of vertices with the same value val_1. Formally,

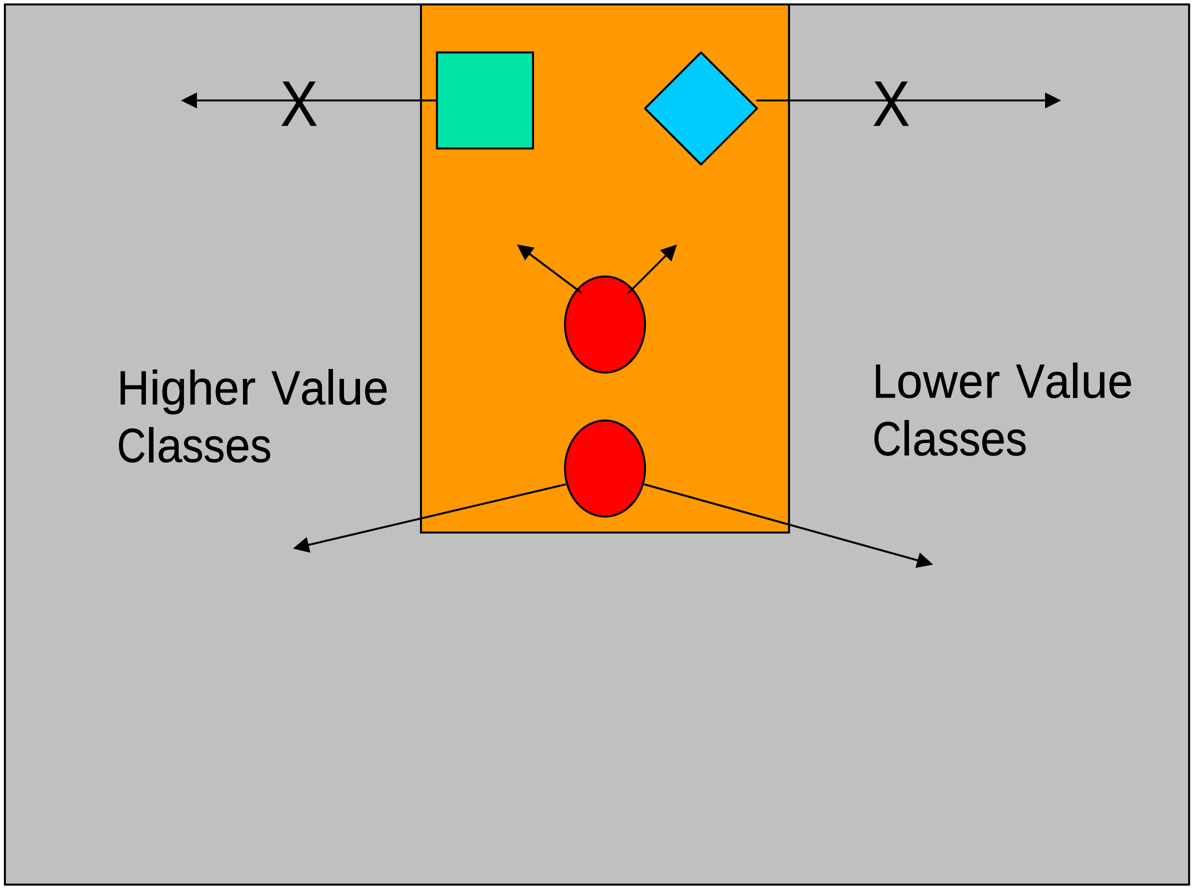
  $C(p)=\{ v : val_1(\Psi_1)(v) = p\}$

- We now observe some structural property of a value class.

Value Class



Boundary Probabilistic Vertices



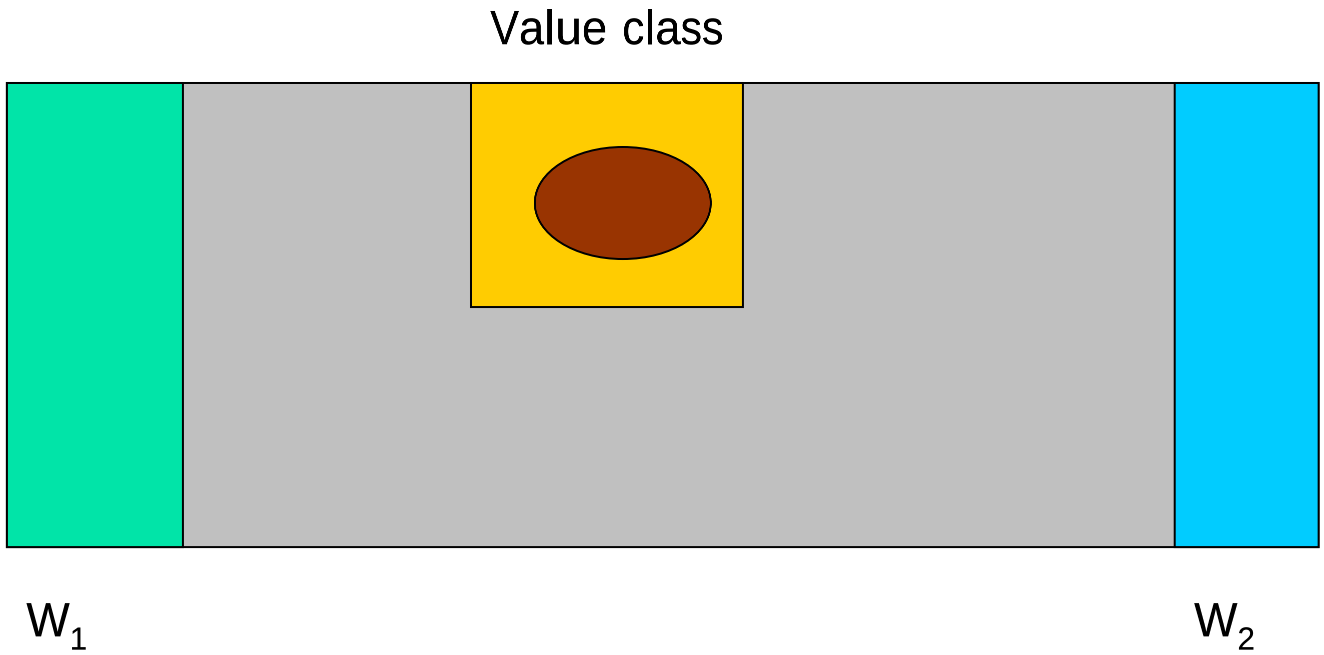
Almost-sure in Value Class

- Sub-game with all edges to other value classes removed.
- Boundary probabilistic vertices changed to winning vertices for player 1.

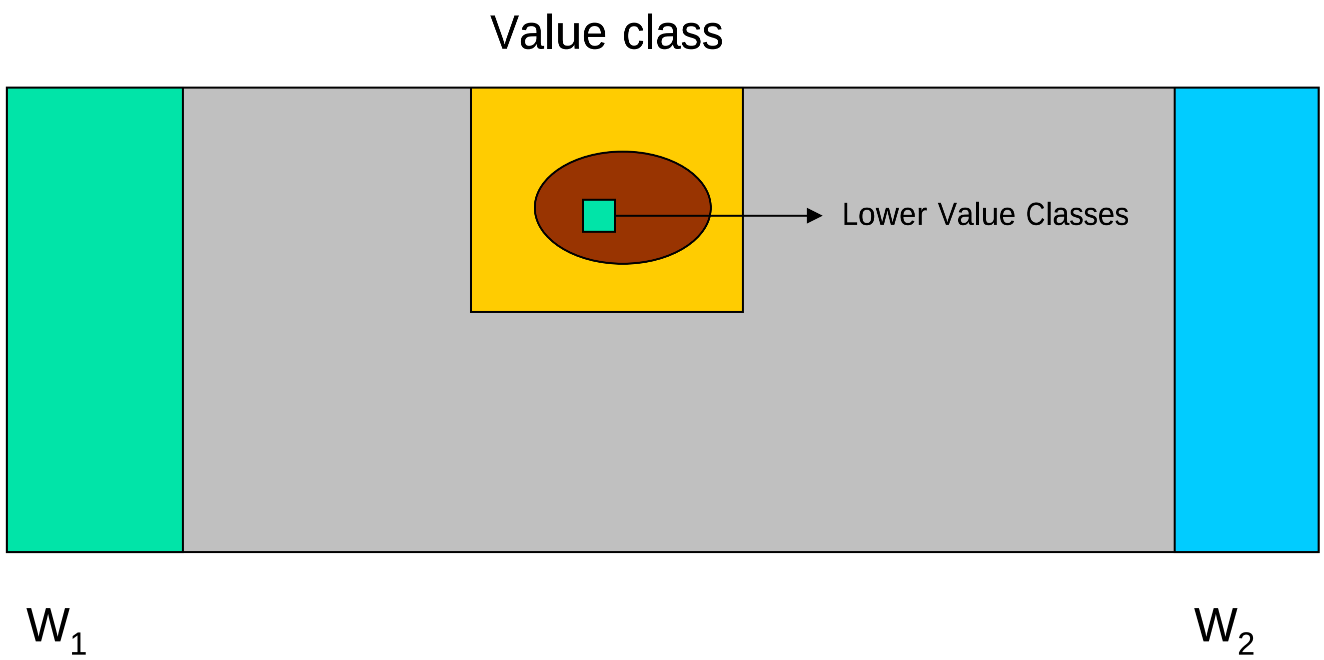
Sub-game Property

- Claim: Player 1 wins almost-surely.
- Proof: Suppose not.
  - Then player 2 wins with positive probability somewhere.
  - Player 2 wins almost-surely somewhere.
  - Player 1 if stays in the value class loses with probability 1 or else jumps to a lower value class.
  - Contradiction.

Almost-sure to Optimal



Almost-sure to Optimal



Value Class Property

- In value classes if we assume boundary probabilistic vertices winning for player 1 then player 1 wins almost surely.
- Conditional almost-sure winning strategies.
- Compose them to get a optimal strategy.

Proof Idea

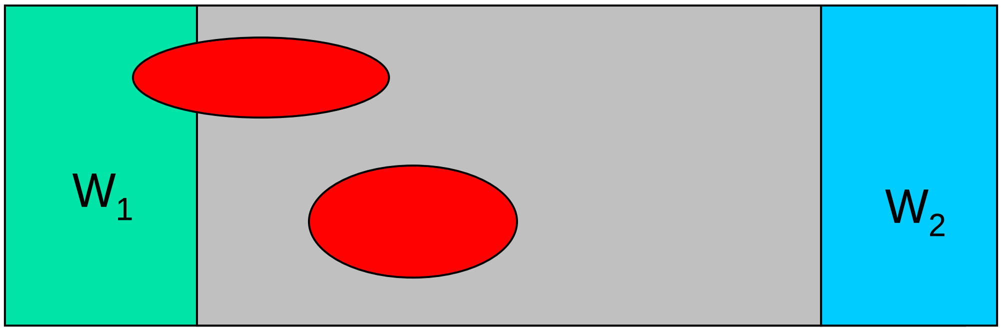
- If the game stays in some value class player 1 wins with probability 1.
- Else it leaves the value class through the boundary probabilistic vertex or goes to a higher value class.
- Invoke sub-martingale Theorem.

Graph Theoretic Argument

- Fix a memoryless conditional almost-sure winning strategy for player 1.
- MDP for player 2: Mueller objective for player 2.
- Fix an memoryless optimal strategy for player 2.

Graph Theoretic Argument

- Markov chain fixing the memoryless strategy for both players.
- Suppose any closed recurrent class touches the gray zone.
- Since conditional a-s strategy for player 1, winning for player 1.
- Since value positive in gray zone, this is not possible.



Graph Theoretic Argument

- Hence reaches $W_1 \cup W_2$ with probability 1. In W_1 value for player 2 is 0.
- Hence reachability to W_2 for player 2.
- Recall player 2 no edge to a better value class and player 1 chooses successor in same value class.

MDPs

- Reachability objectives with target T can be solved by the following linear program.
  - min $\sum_{v \in V} x_v$ subject to
    - $x_v \geq x_w$ $\quad v \in V_1, (v,w) \in E$
    - $x_v = \sum_{(v,w) \in E} x_w \, \delta(v,w)$ $\quad v \in V_0$
    - $x_v = 1$ $\quad v \in T$

Graph Theoretic Argument

- Hence reaches $W_1 \cup W_2$ with probability 1. In W_1 value for player 2 is 0.
- Hence reachability to W_2 for player 2.
- Recall player 2 no edge to a better value class and player 1 chooses successor in same value class.

- Hence assigning $x_v = 1 - val_1(\psi_1)(v)$ satisfies all the constraints.
  - Since the linear program is minimizing player 2 cannot get value greater than its value given the strategy for player 1.

- Thus the fixed strategy for player 1 is optimal.

Basic Methods

- Method 1: Basic requirements of reduction to 2 player games for qualitative analysis.
  - Rabin objectives: pure memoryless almost-sure winning strategy exists.

- Method 2: Construction of optimal strategies from almost-winning strategies in sub-games.

Results

- Follows for all Mueller objectives.
  - Optimal strategies are no more complex than almost-sure winning strategies.
  - In future analysis almost-sure strategy complexity is enough.
- Theorem : 2 ½ player Rabin games have pure memoryless optimal strategies
  - 2 ½ player Rabin-games NP-complete.
  - 2 ½ player Streett-game coNP-complete.
  - 2 ½ player Rabin-chain games NP $\cap$ coNP.

Applications

Algorithms

- Method 1: Reduction to 2 player games for qualitative analysis.
- Method 2: from qualitative analysis to quantitative analysis.
- Also the second reduction is closely related to reachability objectives.

Algorithms

- The previous methods can be suitably used to obtain strategy improvement algorithms for 2 ½ player Rabin games [CH06].
  - Combining the strategy improvement algorithm of Condon for reachability objectives.
  - The improvement step requires solving two player Rabin games.
  - The algorithm can use any two player Rabin game algorithm.

Conclusion

Conclusion

- Conclusion
  - Complexity of Simple Stochastic Rabin and Streett games.
  - The simplest class of strategies suffice for optimality for Rabin objectives in 2 ½ player games.
  - Algorithms such as strategy improvement algorithms.

- Open problems
  - (Conjecture) If pure memoryless optimal strategies exist for 2 player games, then also for 2 ½ player games.
  - Polynomial time algorithms for
    - 2 player parity games.
    - 2 ½ player reachability games.
    - 2 ½ player parity games.

Thanks !!!

http://www-cad.eecs.berkeley.edu/~c_krish